2010-06

Incorporating ensemble-based probabilistic forecasts into a campaign simulation in the Weather Impact Assessment Tool (WIAT)

Palmer, Jeffrey M.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/5256

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

INCORPORATING ENSEMBLE-BASED PROBABILISTIC FORECASTS INTO A CAMPAIGN SIMULATION IN THE WEATHER IMPACT ASSESSMENT TOOL (WIAT)

by

Jeffrey Michael Palmer

June 2010

Thesis Advisor:	Rebecca Stone
Second Reader:	Philip Durkee

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2010	**3. REPORT TYPE AND DATES COVERED** Master's Thesis

4. TITLE AND SUBTITLE Incorporating Ensemble-based Probabilistic Forecasts into a Campaign Simulation in the Weather Impact Assessment Tool (WIAT)	**5. FUNDING NUMBERS**
6. AUTHOR(S) Jeffrey Michael Palmer	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ________________.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited	**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

Stochastic (probabilistic) forecasting techniques give forecasters a means to transmit information about certainty and a range of forecast possibilities to operational decision makers. Previous studies have shown value in probabilistic forecasting through series of independent hypothetical events, or through comparative period forecasts. This thesis demonstrates the value of stochastic forecasting through a series of operational events, in the context of a Strike Warfare campaign in the Weather Impact Assessment Tool (WIAT), a campaign simulator.

Simulated ensemble members and probability files were created to study six weather parameters and their impacts on various Strike Warfare missions. Tests were run comparing deterministic and stochastic forecasts, incorporating varying levels of forecast error and sampled probability thresholds. Metrics within and external to WIAT were employed to analyze the results of the forecasting strategies. Constraints in WIAT's structure and campaign modeling yielded results that suggest, but do not definitively prove, enhanced campaign performance as a result of incorporating probabilistic forecasting. Programming adjustments to WIAT are recommended in order to provide a higher-quality test bed for future studies of both deterministic and stochastic forecasting techniques.

14. SUBJECT TERMS stochastic forecasting, probabilistic forecasting, operational simulation	**15. NUMBER OF PAGES** 137
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

INCORPORATING ENSEMBLE-BASED PROBABILISTIC FORECASTS INTO A CAMPAIGN SIMULATION IN THE WEATHER IMPACT ASSESSMENT TOOL (WIAT)

Jeffrey M. Palmer
Lieutenant, United States Navy
B.S., United States Naval Academy, 2001

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY AND PHYSICAL OCEANOGRAPHY

from the

NAVAL POSTGRADUATE SCHOOL
June 2010

Author: Jeffrey Michael Palmer

Approved by: Rebecca Stone
Thesis Advisor

Philip Durkee
Second Reader

Philip Durkee
Chairman, Department of Meteorology

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Stochastic (probabilistic) forecasting techniques give forecasters a means to transmit information about certainty and a range of forecast possibilities to operational decision makers. Previous studies have shown value in probabilistic forecasting through series of independent hypothetical events, or through comparative period forecasts. This thesis demonstrates the value of stochastic forecasting through a series of operational events, in the context of a Strike Warfare campaign in the Weather Impact Assessment Tool (WIAT), a campaign simulator.

Simulated ensemble members and probability files were created to study six weather parameters and their impacts on various Strike Warfare missions. Tests were run comparing deterministic and stochastic forecasts, incorporating varying levels of forecast error and sampled probability thresholds. Metrics within and external to WIAT were employed to analyze the results of the forecasting strategies. Constraints in WIAT's structure and campaign modeling yielded results that suggest, but do not definitively prove, enhanced campaign performance as a result of incorporating probabilistic forecasting. Programming adjustments to WIAT are recommended in order to provide a higher-quality test bed for future studies of both deterministic and stochastic forecasting techniques.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AOR	Area of Responsibility
ASW	Anti-Submarine Warfare
ATO	Air Tasking Order
BDA	Battle Damage Assessment
CAS	Close Air Support
CAW	Carrier Air Wing
CNMOC	Commander, Naval Meteorology and Oceanography Command
CVN	*USS Nimitz* -class aircraft carrier
DCF	Diagnostic Cloud Forecast
DEAD	Destruction of Enemy Air Defense
DoD	United States Department of Defense
EPS	Ensemble Prediction Systems
ESG	Air Force Environmental Scenario Generator
ESRL	Earth System Research Laboratory
FACs	Field Air Commands
FAC	Forecast Accuracy
FAR	False Alarm Rate
FNMOC	Fleet Numerical Meteorology and Oceanography Center
GCAM-CTS	General Campaign Analysis Model – Core Tool Suite
GUI	Graphical User Interface
hPa	Hectopascals
IC	Initial Condition
IRG	Iraqi Republican Guard
ISR	Intelligence, Surveillance, Reconnaissance
IT	Information Technology
JPITL	Joint Prioritized Integrated Target List
KI	Killbox Interdiction
METOC	Meteorology and Oceanography
MM5	Fifth-Generation National Center for Atmospheric Research / Pennsylvania State University Mesoscale Model

MOE	Measures of Effectiveness
MRF	Medium Range Forecast
NCAR	National Center for Atmospheric Research
NCEP	National Center for Environmental Prediction
NOGAPS	Navy Operational Global Atmospheric Prediction System
NWP	Numerical Weather Prediction
NSW	Naval Special Warfare
OIF	Operation IRAQI FREEDOM
ORM	Operational Risk Management
P_d	Probability of Detection (aircraft sensors)
PDF	Probability Distribution Function
P_k	Probability of Kill
POD	Probability of Detection (weather events)
RMSE	Root Mean Square Error
ROI	Return on Investment
RUC	Rapid Update Cycle
STK	Strike
STW	Strike Warfare
UAV	Unmanned Aerial Vehicle
VBA	Visual Basic for Applications
WIAT	Weather Impact Assessment Tool
WIAT*	Weather Impact Assessment Tool (used for stochastic analysis)

ACKNOWLEDGMENTS

Completing a thesis requires the input and assistance of a large number of people, far too numerous to fully name here. First, I would like to thank Professor Tom Murphree, who helped me get started in this study of WIAT, and F. Anthony Eckel, MAJ, USAF (Ret.), who introduced me to the intricacies of ensemble forecasting.

I greatly appreciate CDR Rebecca Stone for accepting me as a thesis student. I will be eternally grateful for her patience, insight, and guidance, on both professional and scholarly levels. She continually found time for me over the closing months of this process, despite an already full plate of work from her departmental and professorial duties. I am truly fortunate to have had the opportunity to work with her.

I also thank Luke Piepkorn, of Systems Planning & Analysis (SPA), for fielding a constant barrage of questions on programming and theory. Many times Luke made space in his work schedule (and home life) to help me understand the complex simulation that is WIAT, as well as provide IT support when necessary. As with CDR Stone, this thesis would not have been possible without his help.

Completing a technical thesis away from the Naval Postgraduate School is a stressful and daunting undertaking, and it would be truly impossible were it not for operational leaders who understood the issues, and gave me opportunities to find solutions. I will be eternally grateful to CDR David McDowell, CDR R. Mark Sampson, and CDR A.J. Reiss for giving me the shepherding that I needed, when I needed it.

Without the support of my loving and long-suffering wife, Jill, I truly would never have found my way through this wilderness. She endured the ups and downs, extended grace when none was deserved, and found a way to come alongside at so many times when I needed the reassurance and encouragement—her love is amazing.

Thanks also to the legions of friends, family, and colleagues who constantly prayed for me—I can never thank you enough.

Last, but most, I am thankful to God for His love and discipline, both of which I saw continuously during this process.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The 2008 Commander's Guidance from Commander, Naval Meteorology and Oceanography Command (CNMOC) indicates that increased budget and research resources will be devoted to the continued development of probabilistic forecasting techniques and products. Deterministic forecasts provide single values without indication of uncertainty or reliability, while probabilistic forecasts can help improve decision making and combat capability by indicating ranges of possibilities (Eckel 2008). Prior research demonstrates the advantages of stochastic forecasting in the context of a series of repeated, independent decisions (Thompson 1952, Eckel 2008). The focus of this study is to demonstrate the net benefits of stochastic forecast-based decision making in a complex, multi-step weather impacts simulator of Operation IRAQI FREEDOM (OIF), the Weather Impact Assessment Tool (WIAT).

The main objectives are to: 1) demonstrate improvement to simulated campaign operations by utilizing probabilistic forecasting, and 2) investigate the existence of optimal decision probability thresholds for weather parameters in the simulated campaign. The main benefit of this study will be the application of stochastic forecasting to operational risk management (ORM) in the context of complex operational decision making in a simulated strike warfare campaign. It will also provide an example of how probabilistic forecasting can be incorporated into a variety of similar warfare types beyond strike warfare.

This thesis is organized into five chapters. In Chapter II, we discuss the rationale for developing stochastic forecasts. Chapter III describes the WIAT program. Chapter IV presents the methodology for running WIAT in deterministic and stochastic modes to meet the above objectives. Chapter V shows output metrics and compares the results of campaigns run under deterministic and stochastic frameworks. Conclusions and recommendations are presented in Chapter VI.

THIS PAGE INTENTIONALLY LEFT BLANK

II. BACKGROUND

A. SCOPE

This chapter will discuss ensemble forecasting and how operational risk management relates to forecasting. An explanation of the sources of error in numerical weather prediction (NWP), deterministic and stochastic prediction methods, and characteristics of ensemble prediction systems (EPS) is fundamental to understanding the benefit of including forecast uncertainty information in the decision process. Proper application of ensemble forecasts requires an assessment of costs, risks, and rewards associated with employing certain units or executing missions.

B. ATMOSPHERIC PREDICTION

1. Uncertainty

Lorenz (1963) found that small adjustments to the initial state of a dynamical system led to considerably different results as time progressed. Based on this experimentation, he suggested that error in describing the current state of the atmosphere is the major contributor to difficulties in predictability. This analysis of the atmospheric state is also known as the atmosphere's initial conditions (ICs), a description of the values associated with the state of the atmosphere at a given point in time. However, the analysis ICs will never be precisely the same as the actual atmospheric ICs. Error in determining analysis ICs can result from imperfect measurements, from the sparse and uneven distribution of atmospheric observations, and from assumptions and approximations used in assimilating the observations. Limited model resolution also adversely affects data assimilation, and in turn, the ICs (Kalnay 2004).

The earth's atmosphere is a dynamical system, and accordingly should be predictable provided that its current state and the rules for its evolution can be completely deduced. After determining the ICs for the atmosphere, operational models employ a set of equations that approximate relationships among various atmospheric parameters. However, there are several pathways for error to be introduced into the model. The differential equations that govern the development of atmospheric motion over time do

not completely describe all motion within the atmosphere. Difficulties in modeling both upper and lower boundaries of the atmosphere, approximations made in the numerical solution techniques, various simplifications of small-scale processes, and finite computer precision all contribute error, so that numerical models cannot exactly project future states of the atmosphere (Eckel 2008).

Because both the analysis and the model are flawed, error introduced by the analysis is magnified by forecast error, and the difference between the actual state and the model state grows with increasing lead time. Even the finest model resolution available will not prevent the compounding of small errors over time, and the corresponding divergence of modeled weather from true weather.

As a result of the increasing divergence of true weather and modeled weather, prediction capability diminishes with increasing lead time. The limit of predictability for a specific type of forecast (a given atmospheric variable) occurs at the lead time where the mean squared error of the model forecast equals the mean squared error of a forecast of the climatological mean (Eckel 2008). Beyond this lead time, the operator receives better forecast information from climatology than a model forecast.

2. Deterministic and Stochastic Prediction Methods

The method most commonly employed for atmospheric weather prediction is the deterministic method, characterized by a single-value forecast or descriptive solution. The great strides made in both computing power and model refinement over the last few decades have extended predictability, but even so, short- and medium-term forecast errors are occasionally large due to one or more of the sources of error discussed previously. One example of such a scenario was the over-forecasted Thanksgiving Storm of 2001. The 24-hour forecast from the Fifth-Generation National Center for Atmospheric Research / Penn State University Mesoscale Model (MM5) for 22 November 2001 (Figure 1) depicted an operational forecast for gale force winds in the Puget Sound and Straits of Juan de Fuca (Eckel 2008). Observations of that period, however, indicated calm weather. Later analysis of the model, shown in Figure 2, indicated that the operational forecast had depicted a plausible but low-probability event, given multiple re-runs of the model; in other words, the model had forecast a possible

event, but the forecasters had no indication of the low likelihood of this forecast. Ultimately, deterministic forecasts greatly limit the weather forecaster's ability to convey valuable information about uncertainty and multiple possible events. This leads to weather forecast products of both numeric and graphical types that convey a false sense of confidence to the user, which negatively affects the quality of decision making by operational users.

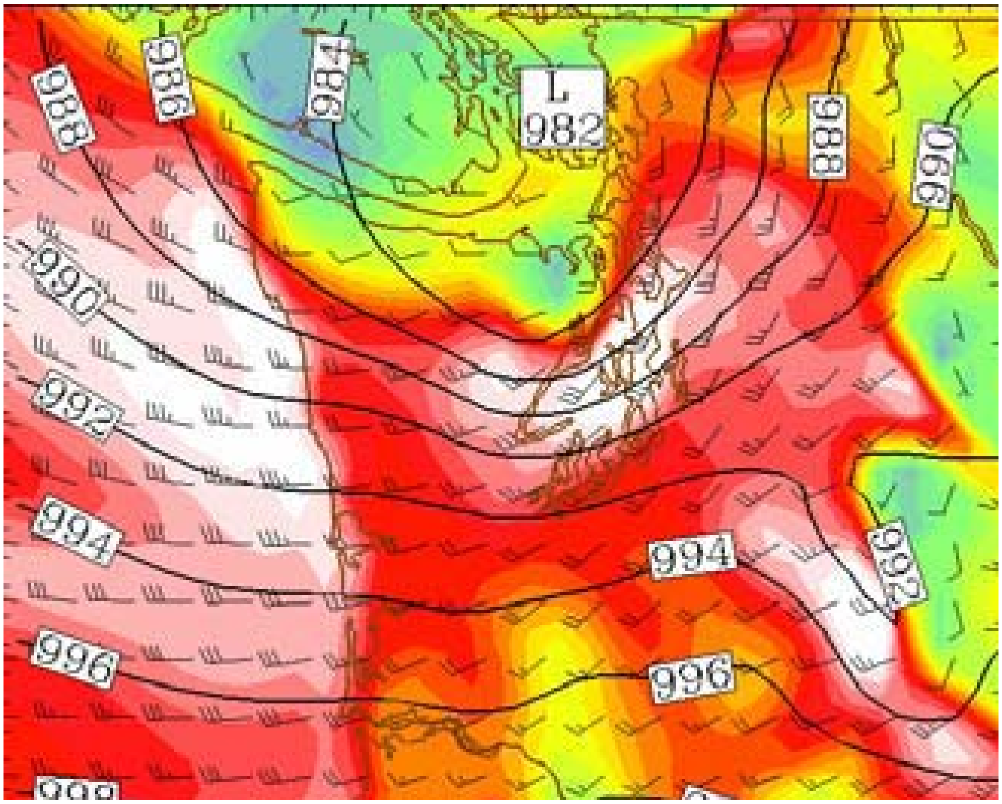

Figure 1. MM5 Operational 24-hour Forecast, depicting a gale force wind event in the Puget Sound and Straits of Juan de Fuca. Valid 1800Z, 22 November 2001. From (Eckel, 2008).

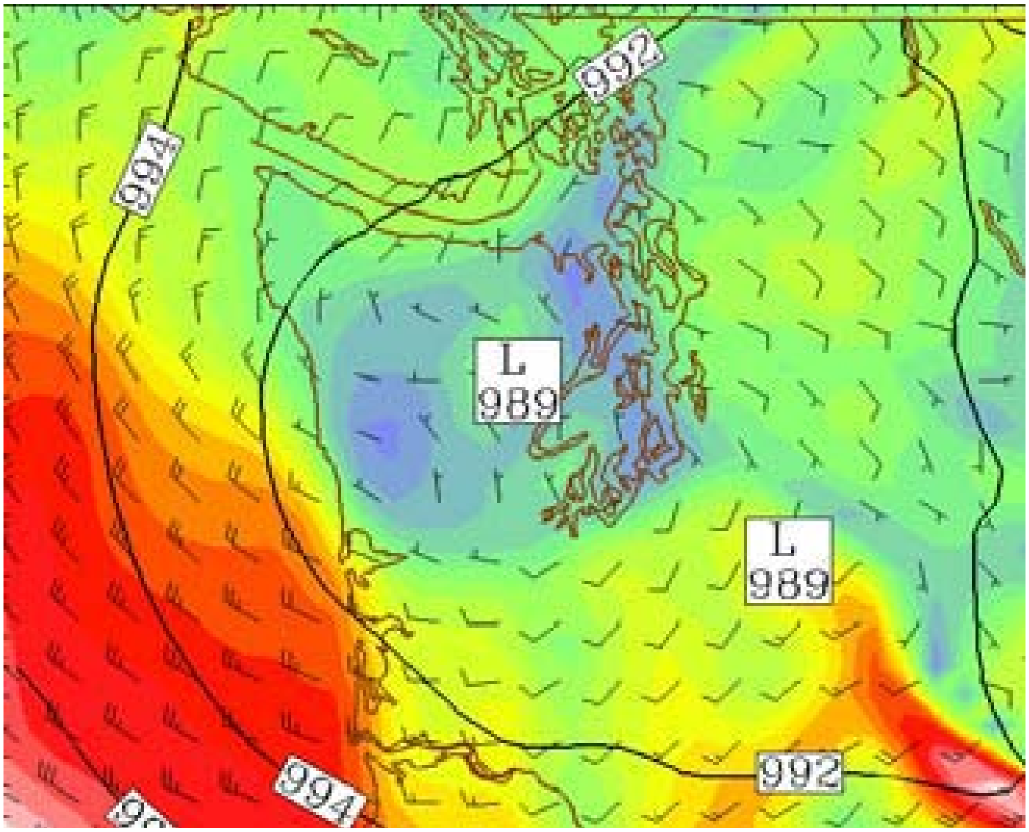

Figure 2. MM5 Verification Analysis for 1800Z, 22 November 2001, depicting a weaker low-pressure area than forecast, with no gale force winds. From (Eckel, 2008).

Stochastic forecasting, on the other hand, provides a suite of solutions that objectively convey potential outcomes and forecast uncertainty. Such information allows the user to grasp the range of weather conditions (evolved from the present state) that may impact his operations. If the user has constraints in operations that would be broken by one or several potential outcomes, a stochastic approach permits a better decision-making process by providing a more complete set of information. Instead of focusing solely on limiting the forecast error, stochastic forecasting attempts to model the flow-dependent error growth, and then depict the uncertainty in a given forecast (Eckel 2008).

C. ENSEMBLE FORECASTING

1. Description

While there is no standard means of deriving an ensemble forecast, the most general method involves repeatedly adding small perturbations to both analysis and model to produce multiple unique forecast solutions. Zhu et al. (2001) state that the operational ensemble forecasts produced by multiple prediction centers around the world use “models…integrated a number of times, started from slightly perturbed initial conditions, in addition to generating the traditional ‘control’ forecast that starts from the best available atmospheric analysis.” This incorporation of slightly perturbed initial conditions (ICs) simulates the analysis error inherent in portraying the atmosphere. Each different IC may then be funneled into a slightly different version of the given NWP model, as a means of accounting for model error. Systems that use only one version of the model (an unperturbed model) effectively neglect the model as a source of error and uncertainty in the ensemble. Taken all together, the ensemble members are used to represent a spectrum of potential future states of the atmosphere. This spectrum provides estimates of forecast uncertainty and probabilities for specific events. Each member should have an equal chance of being the true outcome; incorporating more ensemble members (model runs) gives the ensemble user a better chance of at least one member being close to the truth.

One of the meteorological prediction centers responsible for producing ensemble forecasts is the Fleet Numerical Meteorology and Oceanography Center (FNMOC), in Monterey, California. FNMOC produces a number of ensemble products based on the Navy Operational Global Atmospheric Prediction System (NOGAPS) model, an example of which is shown in Figure 3. The ensemble products, instead of displaying contours of atmospheric parameters such as pressure, winds, and temperature, as shown in the MM5 graphics previously, display contours of probabilities of exceeding a given meteorological threshold. In the case of Figure 3, the contours are of the likelihood of exceeding 6.25 mm precipitation in the twelve-hour period ending at 0000Z 13 September 2008.

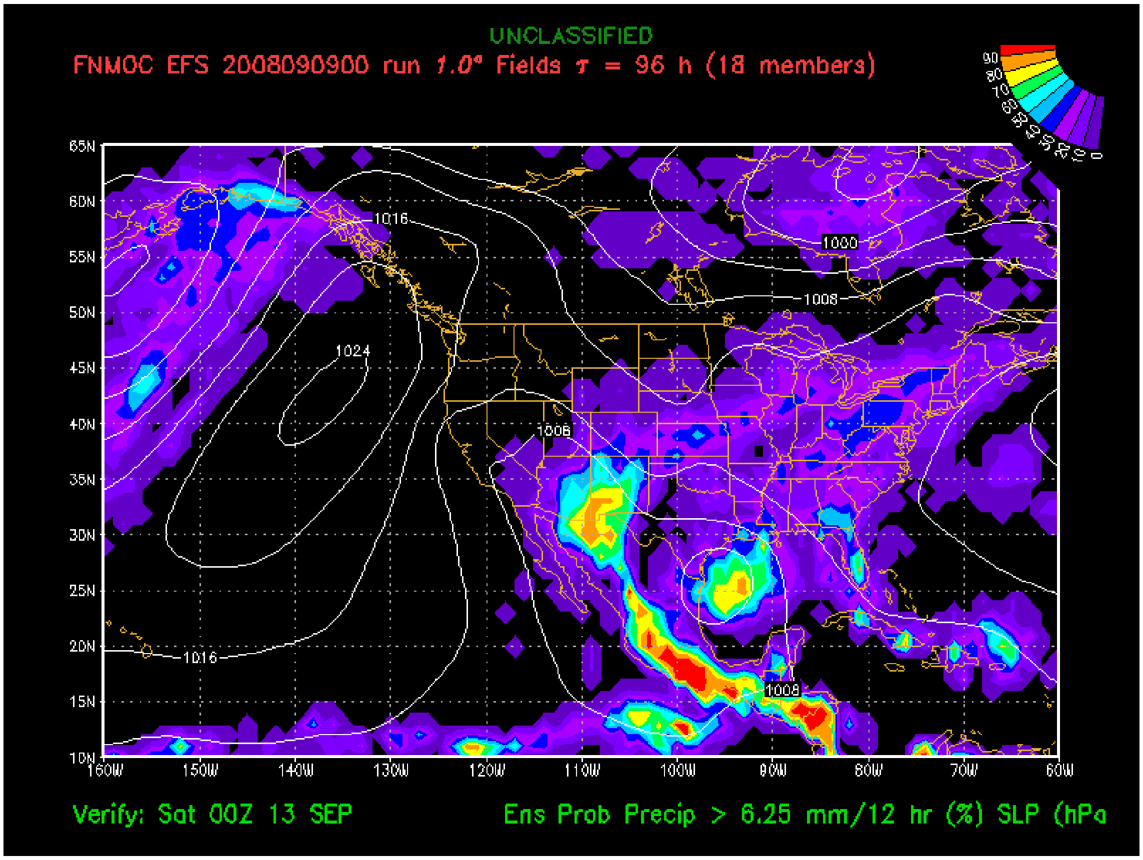

Figure 3. FNMOC EFS Graphic displaying 96-hour Ensemble Forecast sea level pressure (contours) and probabilities of precipitation exceeding 6.25 mm over the period from 1200Z 12 September 2008 to 0000Z 13 September 2008. From FNMOC Web site.

Both deterministic and stochastic techniques employ atmospheric observations, analyses, modeling, and forecasts, in the effort to describe the atmosphere. Stochastic forecasts, however, portray information as probabilities or as probability density functions (PDFs). Ensemble forecasting pursues the production of a "sharp, calibrated forecast probability density function (PDF) of possible future states of the atmosphere from which the true state is consistently a random sample" (Eckel 2008).

The evolution of the atmosphere is fundamentally a deterministic process. In principle, there is an exact set of characteristics that describe the atmosphere at any given time, and a set of rules that comprehensively governs the evolution of the atmosphere towards one specific state at a given future time. In practice, this analysis and evolution cannot be known exactly, but we can construct a system that will provide a good sense of

the likely analyses and changes over time. A likely state of the atmosphere could be represented as a point in a PDF that includes all possible states of the atmosphere at that time (a true analysis PDF). A possible future state of the atmosphere could also be represented as a point on a PDF that includes all potential future states (a true forecast PDF.) Hence, the ensemble forecast process pursues the production of an analysis PDF that closely depicts the true analysis PDF, and a forecast PDF that approximates the true forecast PDF. Neither the analysis PDF nor the forecast PDF are absolute, as they are dependent on the forecast system used to derive them; for example, if more observations are incorporated into the forecast system, the shape of the analysis and forecast PDFs will change. In other words, the ensemble forecast process attempts to approximate a hypothetical distribution of weather states from which the truth is always a random sample.

2. Calibration and Ensemble Size

Several problems are inherent to the ensemble's forecast PDF. In an ideal ensemble, the slightly perturbed analyses and models incorporate all sources of uncertainty; in reality, ensembles can only include the primary sources of uncertainty, and only to a limited degree. Such limitations lead to forecast PDFs, which differ significantly from the true forecast PDFs, and thus forecast probabilities (interpreted as forecast frequencies of occurrence) do not agree with observed frequencies of occurrence. These forecast PDFs are said to be uncalibrated or poorly calibrated ensemble forecasts. A calibrated forecast has been adjusted so that the estimated probability of an event matches the rate of occurrence of the event, i.e., the forecast PDF is adjusted so that it approaches the true forecast PDF. For example, an ensemble initiated with a given meteorological state returns a 40% probability of an event occurring at some point in the future. We should expect that if the same meteorological situation occurs 100 times in nature, the forecast event would occur 40 times. If instead, the event occurred 30 times, it becomes apparent that the method used in forecasting the event overestimates the probability of occurrence. Even if we cannot find and correct the cause of the overestimate, we can calibrate by adjusting the 40% forecast to 30%.

Typically, narrower PDFs indicate confidence in a smaller range of potential outcomes, while wider forecast PDFs indicate more uncertainty in the forecast. For example, a forecaster is interested in predicting 500-hectopascal (hPa) geopotential heights for a given location. Provided that a good analysis is made, the forecaster can expect to see a fairly narrow range of potential outcomes for a 12-hour forecast, a relatively short lead time. With increasing lead times, however, there is greater opportunity for analysis errors and model errors to interact and compound each other; for a 48-hour forecast of 500-hPa geopotential heights, the forecaster would expect to see a larger range of values as potential solutions. This reflects the increased uncertainty with increasing lead time, and accordingly wider PDFs.

Finally, an ideal ensemble would incorporate an infinite number of members that could "fill out" the forecast PDF. With an infinite number of members, all potential forecast states would be represented, and a complete distribution could be constructed. Instead, EPSs sample the forecast PDF a finite number of times. Incomplete sampling of the forecast PDF introduces an additional source of error, in that the EPS will not accurately depict the range of values in a multi-dimensional PDF (pockets of values may be over- or under-represented.) In order to achieve forecast PDFs that approximate the true forecast PDFs, most EPSs incorporate 20–30 model runs to ensure a skilled forecast, with upwards of 50 runs if the system is particularly designed to capture low probability events found at the tails of the forecast PDF. (Eckel 2008).

D. COST-LOSS AND VALUE ANALYSIS

Brier (1944) was among the first to suggest using probabilities of given classes of weather events in creating a forecast instead of simply forecasting one particular class of event. Brier used an example of a movie producer's concern with stratus cover during a photographic shoot as a means to introduce a cost-loss scenario. In his example, the producer can complete his work if the skies are clear after 1:00 PM, and the most likely forecast time for cloud dissipation is 11:30 AM, with some possibility that this forecast may be wrong by several hours. The producer is then interested in the probability, P, of this cloud dissipation before 1:00 PM. The cost to the producer to hire help, set up equipment, etc. is the stake (S) put up to achieve a gain (G) realized by completing the

work that day. If this decision were repeated over multiple trials ("in the long run"), the producer could expect to accumulate a net benefit if $P > S/G$. This means that the producer should only risk the cost associated with setting up the shooting event when the probability of clear skies exceeds the ratio of the stake to the prospective gain from holding the event.

In Brier's example, the stake is $50,000 and the gain is $100,000. The probability assigned to the chance of clear skies is 0.8. If ten trials were held, and the forecast probability was calibrated (outcomes matched the probabilities), then the producer would see a total stake expenditure of $500,000 (10 trials multiplied by the $50,000 stake per trial) and a gain of $800,000 (eight trials where the sky was clear multiplied by the $100,000 gain per successful trial), for a net benefit of $300,000. However, if the probability assigned to the chance of clear skies is calibrated at 0.4, the producer could expect to see a negative return over the series of decisions (for ten trials, -$100,000.) Both of these scenarios assume that the producer will put forward the stake in every trial. However, if the producer understands that the chance of clear skies is less than the stake-gain ratio (in this case, 0.5), then he will decide not to expend setup money for any of the trials, and accept a net benefit of 0 rather than the net loss of $100,000. The stake-gain ratio becomes the producer's break-even point and consequently the decision threshold for conducting the shoot. This sort of threshold is integral to the campaign decision-making process both in this study's use of WIAT and in real-world scenarios, as it can be utilized to evaluate the net benefit of conducting an operation within a campaign, as well as finding the threshold where benefit is optimized over a longer time frame.

Thompson (1950) used a hypothetical case study to demonstrate the ability to minimize monetary costs taken for protective measures by using a cost-loss comparison with the probability of a given meteorological event. His posited construction company was responsible for pouring concrete in Los Angeles over the October 1949–March 1950 period. He assigned a cost of $400 to protect concrete from a precipitation event, while the loss associated with not protecting the concrete (given the occurrence of 0.15 inch of

rain within 36 hours of the time the concrete was poured) of $5000. He conducted a comparison of the resulting expenditures according to six different strategies that could have been undertaken by the company:

(1) receiving no forecast at all,

(2) receiving no forecast but taking protective measures daily,

(3) employing protective measures only when climatological expectancy exceeded the cost-loss ratio (equal to 0.08),

(4) using persistence (employing protective measures on all days following a day with measurable precipitation),

(5) using a categorical forecast based on whether the probability of the rain event was greater or less than 50%,

(6) or using a forecast based on whether or not the probability of the rain event exceeded the cost-loss ratio (Thompson 1950).

The last of these options was found to yield the lowest total expenditure, a full 33% less than the categorical forecast.

The case studies presented by Thompson and Brier effectively illustrate two different means of using probabilities to improve outcomes for the user. Brier's movie producer is working in a cost-gain framework, where the object is to maximize benefit (in this case, return on investment in the form of completed photographic shoots.) Thompson's construction company is operating in a cost-loss framework, where the object is to minimize loss (payment to provide protective measures in case of thunderstorms.) Both case studies, however, demonstrate the balancing act required in accepting an appropriate amount of risk as compared to the user's specific risk tolerance, as a means to optimize long-term performance.

While Brier speculated on a specific either/or scenario used to protect the budget of a movie producer, Thompson expanded the list of options available to a decision maker. Comparing consequences from following various sources of forecast information, the tailored forecast information based on the company's cost-loss ratio would have yielded the lowest expenditure for the company. Thompson interpreted these results to

show that probability estimates had advantages, and that deterministic forecasts contained an "inherent danger" due to the fact that "the user is not provided with, or does not make use of a measure of the reliability of the prediction."

Zhu et al. (2002) evaluated the economic value derived from the use of a 14-member ensemble of the Medium Range Forecast (MRF) T62 model (referring to global spectral model that resolves 62 zonal wave numbers), as compared to a single T62 run and a single T126 run. A forecast has economic value if the user does better using the forecast than using a climatological forecast; the writers' experiment was designed to determine which of the deterministic and stochastic forecasts would most out-perform the climatological forecast. While the authors did not specify a specific real-world cost-loss decision, reducing the cumulative cost of protective measures characterized the discussion, much as it did in Thompson's and Brier's work.

Zhu et al. found that for prediction of 500-hPa heights over Northern Hemisphere extratropics, all users were better off with the ensemble forecasts after 120 hours, with the vast majority better off with ensemble forecasts after 72 hours. The authors found that for cost-loss ratios outside of a band from 0.2 to 0.5, the ensemble forecasts provided better information than either of the control forecasts. They also demonstrated that the range of cost-loss ratios for which forecasts demonstrated value was substantially wider with the use of ensemble forecasts. Since the ensemble of 14 T62 runs represented roughly the same computational cost as the higher-resolution T126 single run, the findings suggest that for certain situations, ensemble runs of a lower-resolution model can be more cost-effective than a higher-resolution model run.

The authors state that ensemble-based distributions can exceed the performance of the single control run partly due to the nonlinear filtering that takes place when forming the ensemble mean, but more due to its ability to reflect time and space (flow-dependent) variations in forecast uncertainty (Zhu et al. 2002). The deterministic forecasts had more economic value than the ensemble at lead times shorter than 72 hours; the increased resolution enabled better characterization of the analysis state and the evolution of the system over short time scales. However, the deterministic forecasts cannot reveal the effects of non-linear error growth, as stochastic forecasts can, and hence deterministic

forecasts become less reliable with increased lead time. Additionally, the ensemble forecasts could portray degrees of uncertainty to users; this capability is lacking in deterministic forecast guidance.

Zhu et al. also reinforce the idea, previously discussed by Thompson (1952), that the probabilistic forecast is more useful in that it allows users to establish their own decision criterion (i.e., a cost-loss ratio.) Instead of being provided with a deterministic forecast that suggests an “either/or” decision, the forecast user can now compare a probability of occurrence with their own specific risk tolerance; savvy users will compare the probabilistic forecast with their own cost-loss estimations to make decisions that will provide value over the long term (Zhu et al. 2002).

Zhu et al.’s and Thompson’s findings have particular applicability to Department of Defense (DoD) weather forecasters, who advise decision makers involved in conducting operations, training, and managing resources. The ability of the forecaster to accurately predict the most likely evolution of weather parameters and to communicate a qualitative level of uncertainty is necessary, but not sufficient, for optimal decision making. The decision maker needs to know about less-likely but high-risk weather events; he or she needs to be able to “calculate the odds.” In many arenas, and often in the armed forces, decision makers expect weather forecasters to give them “go/no-go” recommendations; this leads the forecaster to combine the weather prediction with their own guess about the commander’s risk tolerance, and to do it implicitly rather than explicitly. In an attempt to clarify the concept of what makes a “good” forecast, Murphy (1993) describes three aspects of forecast evaluation: consistency, quality, and value. Consistency (“type 1 goodness”) refers to how well the forecast reflects the forecaster’s best judgment, quality (“type 2 goodness”) refers to how well forecast conditions correspond to observed conditions at or around the forecast valid time, and value (“type 3 goodness”) is determined by whether the forecast yields some degree of increased benefit to the forecast consumer (Murphy 1993). Under Murphy’s construct, many DoD forecasters are expected to make an ex-ante determination of the decision maker’s optimum threshold, often at the expense of consistency. Probabilistic forecasts offer the forecaster an opportunity to better meet the goal of consistency, while also permitting

optimization of value. With full information, including quantitative uncertainty estimates, decision makers can optimize decisions in a cost-loss scenario as long as the cost of protective measures and unprotected loss can be defined. For more complex situations, value must be derived differently.

Eckel et al. (2008) used two operational scenarios to demonstrate the benefit of employing stochastic forecasts in conducting operational risk management (ORM). Previous work detailed above focused exclusively on protection measures, referred to as "defensive ORM." Eckel et al. extended the application of ORM to offensive operations, illustrating advantages from employing ensemble forecasting for gains rather than just for resource protection.

The authors used an evacuation simulation, an example of defensive ORM, to show that over an extended time frame, decisions based on stochastic forecasts will result in fewer cumulative dollars spent for purposes of asset protection. In this case, the stochastic operator yielded a 30% savings rate over the use of deterministic forecasts. In the offensive ORM scenario, stochastic forecasting in for destruction of enemy air defense (DEAD) missions provided a superior operational decision twice as often as the deterministic solution. The stochastic forecasts yielded mission completion 10% faster than their deterministic forecast counterparts. Eckel et al. demonstrated that stochastic forecasts can improve both efficiency and war-fighter effectiveness, in both offensive and defensive ORM scenarios. Net economic value is not defined in the DEAD simulation; the quantified benefit to campaign planners is the earlier destruction of enemy emplacements.

E. DIRECTION OF THIS STUDY

This thesis will follow in the direction suggested by Eckel et al. and Zhu et al. by applying stochastic forecasting to operational decision-making, but extending the context to the more complex campaign scenario in WIAT. It builds upon the work done by Eckel previously, which compared outcomes for a series of independent events. The use of WIAT places the ensemble forecasts in the context of a dynamic environment where each forecast has direct impact on mission planning, the execution of the mission, and hence the necessity of related follow-on missions. Simulation of more complex scenarios,

where use of information from the ensemble forecasting system shows benefit, should help demonstrate to decision makers that information about uncertainty, rather than being a nuisance, has operational value.

III. THE WEATHER IMPACT ASSESSMENT TOOL (WIAT)

A. WIAT AND SIMULATIONS

War games have increasingly become a staple of military planning worldwide over the last two centuries; the American Navy's history of such study extends over a century, beginning with the 1880s and the early years of the Naval War College. Because battles in air, on land, and at sea are so expensive (in terms of money and human life) and relatively unique in comparison to each other, operational planners have long sought means to try out their strategies prior to employing them in actual battle. In the computer era, simulations have become an increasingly large aspect of not only operational planning, but platform and weapon design as well. Simulations permit many iterations of tests and adjustments using computer hardware and software that are considerably less costly than the consequences of a poorly planned fighter acquisition or a faulty asset allocation.

WIAT is a "stochastic simulation toolset used for assessing the impacts of METOC forecasts and METOC phenomena on strike operations" (SPA 2007a). WIAT was developed to help the Oceanographer of the Navy quantitatively link improvement and/or degradation of forecast quality to improvement and/or degradation of warfighter metrics. The program was designed specifically to support analysis of a Strike Warfare (STW) campaign scenario; however, the program's designers noted that its methods could be applied to Naval Special Warfare (NSW) and Anti-Submarine Warfare (ASW) missions.

WIAT is programmed using the General Campaign Analysis Model—Core Tool Suite (GCAM-CTS), which is a "set of object-oriented model development tools that allow the military analyst to quickly and easily create stochastic simulations" (SPA 2007a). Many prior simulation programs focused on one aspect of campaign planning; GCAM-CTS was developed to incorporate system analysis (i.e., performance of a given weapons platform), engagement analysis (actual combat between given military units), and mission analysis (success of the assigned mission) as part of evaluating campaign strengths, weaknesses, and overall effectiveness. The tool suite is able to incorporate

data from a variety of sources and to provide measures of effectiveness (MOEs) that are commonly understood across the various armed forces. GCAM-CTS facilitates repeatability so users can test the sensitivity of various mission and campaign elements to small changes in conditions; WIAT provides an interface that allows the user to run this sort of analysis without any sophisticated programming background. (SPA 2007a)

B. THE CONSTRUCT OF THE WIAT CAMPAIGN

The program simulates a strike warfare (STW) campaign in an operating area covering Syria, Jordan, northern Saudi Arabia, Iraq, Kuwait, and western Iran. The Navy task force in WIAT is composed of a carrier air wing (CAW), which includes both manned aircraft and unmanned aerial vehicles (UAVs), based onboard an aircraft carrier (CVN) operating in the eastern Mediterranean Sea. The wing is tasked with conducting close air support (CAS) operations in support of "ground forces in transit and during engagements," as well as strike (STK) and kill box interdiction (KI) missions with targets exclusively within the borders of Iraq (SPA 2007a). The CAS mission details aircraft to cover friendly "Blue" ground units (brigade-level task forces consisting of armor and/or mechanized infantry units) as they make a simulated move from the northern and southern borders of Iraq towards Baghdad. STK missions have specific assigned targets, such as mobile ballistic missile launchers or troop elements, while KI missions assign aircraft an area inside Iraq to search for a target, and engage it if one is found. Every mission is conducted by two carrier-based aircraft of the same type (i.e., manned and unmanned aircraft do not operate as a mixed sortie.)

Opposing "Red" forces are Iraqi Republican Guard (IRG) divisions, infrastructure elements (i.e., Headquarters units), and mobile land-based ballistic missile launchers. Because the campaign scenario is STW-focused, ground operations are highly simplified; Blue ground forces have a fixed route and engage Red ground forces en route to Baghdad, while Red IRG units attempt to protect Headquarters elements, patrol their immediate areas, and engage Blue forces.

Intelligence, Surveillance, and Reconnaissance (ISR) is provided for Blue forces by both unmanned and manned aircraft, as well as Field Air Commands (FACs) embedded with ground units. These units provide targeting information for CAS and

Strike missions, as well as battle damage assessments (BDA) for previously conducted missions. The UAVs in particular conduct the BDA mission on randomized routes over the area of responsibility (AOR).

C. WEATHER CONDITIONS IN WIAT

Weather conditions in WIAT are modeled in a 0.5° x 0.5° grid, with a varying time resolution (minimum of two hours.) Six weather parameters are included: cloud coverage (upper-level, mid-level, and lower-level) measured in octas, precipitation in millimeters/hour, visibility in meters, and wind speed in meters/second. Weather is read from a Microsoft Excel workbook file, with one worksheet for each time increment of the simulation. Each worksheet has eighteen grids representing a map of the geographic domain: the first set of six grids holds the true weather values for the six weather parameters; the second and third sets of grids are designed for execution and planning forecast values, which can be input by the user or generated by WIAT. Each map is a 20x34 grid, with each cell representing one half-degree square, from 20.508N, 34.474E in the southwest to 30.008N, 49.974E in the northeast. (SPA 2007a).

Data provided to WIAT can come from a variety of sources, but extensive instructions are provided with WIAT if the user would like to obtain data from the Air Force Environmental Scenario Generator (ESG). A default weather file is included with program installation, covering the period January 1–7, 1993.

D. THE OPERATIONAL DECISION-MAKING PROCESS

The program models the planning process for hours to days prior to mission execution, and incorporates weather conditions into operational decision making and impacts through the course of the campaign. Initial flight plan generation starts with the Joint Prioritized Integrated Target List (JPITL) and information from ISR and FACs to produce a Daily Target List and an Air Tasking Order (ATO). This list is used to prioritize sorties and allocate resources (manned and unmanned aircraft) to the various mission types. Once the simulation reaches a point 24 hours before a given mission is scheduled, the planning forecast is checked. If the forecast values exceed the mission thresholds, then the mission is cancelled and replaced; if the forecast is below the

threshold, the mission remains on the task list. Another check is made against the execution forecast six hours prior to mission departure, and if the forecast values are still below mission threshold values. WIAT's inputs for campaign planning can be seen in Figure 4, the method by which planning and execution forecasts is shown in Figure 5, and a sample individual mission with weather impacts is shown in Figure 6 (SPA 2007a).

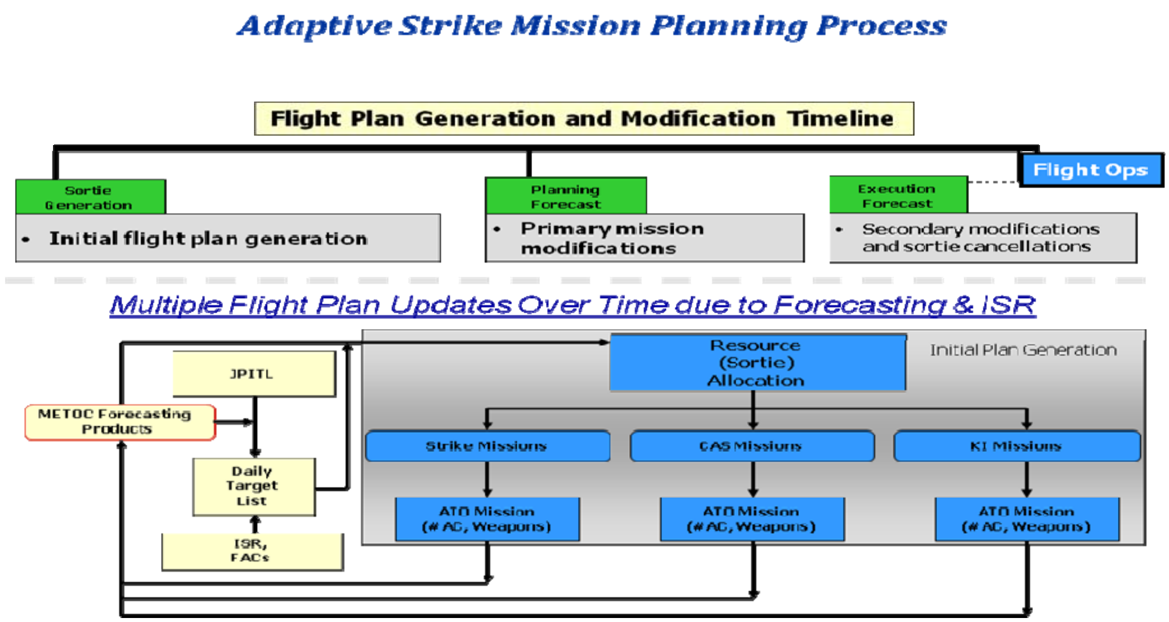

Figure 4. Depiction of Strike Planning Process in WIAT – Flight Plan Generation and Modification (SPA 2007a)

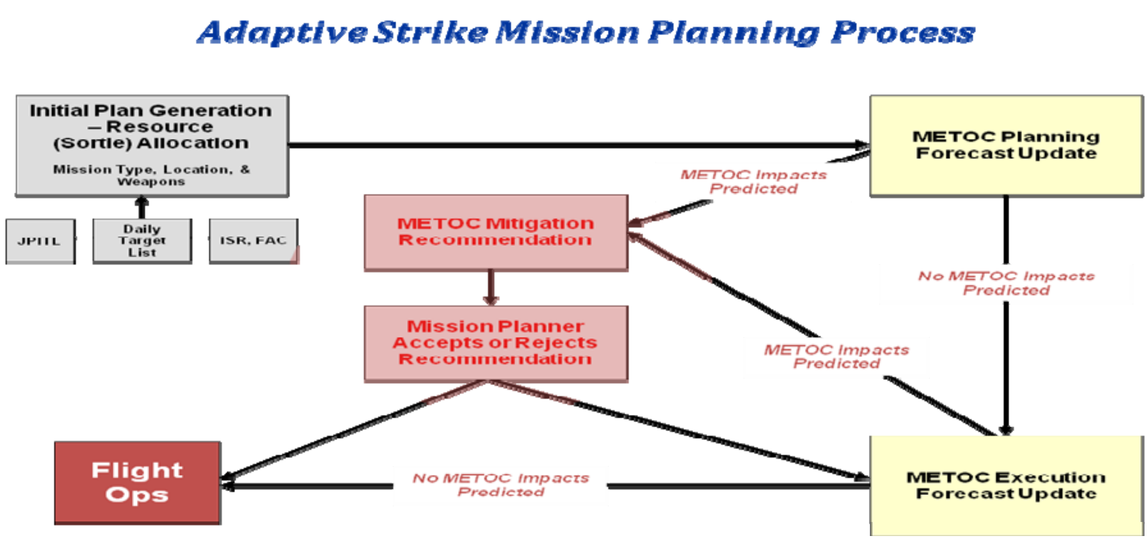

Figure 5. METOC Forecast Integration into Strike Warfare Planning Cycle in WIAT (SPA 2007a)

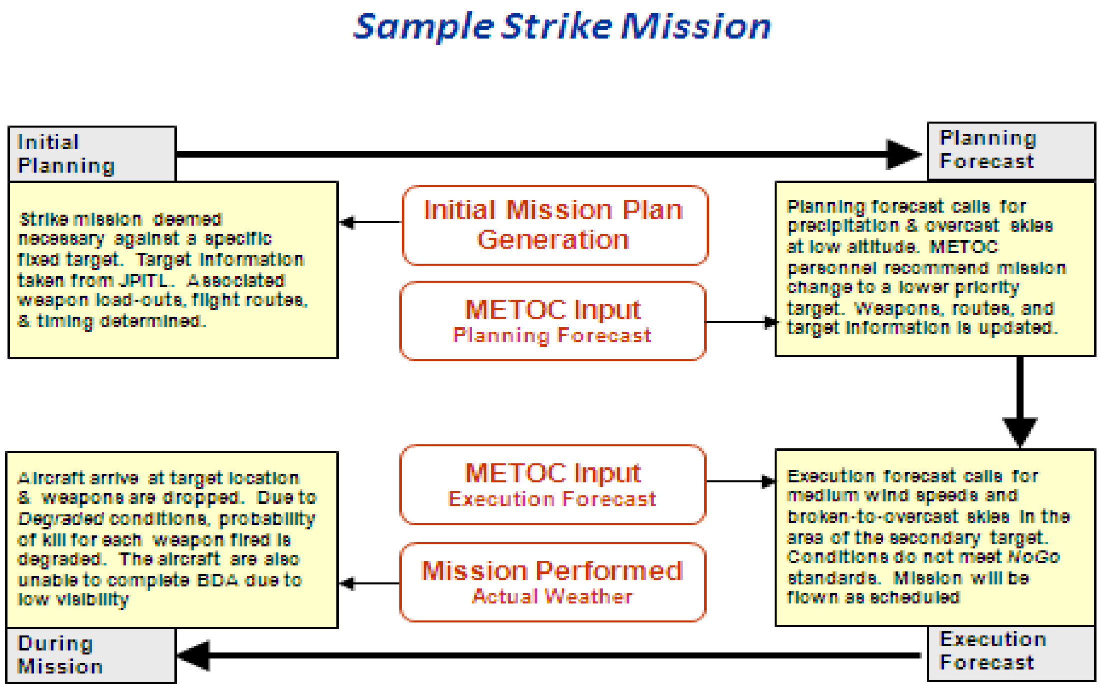

Figure 6. Sample mission progression. After "WIAT Introductory Briefing," 11 June 2007.

E. MISSION EXECUTION

Once the mission departs, the forecaster’s involvement is over; however, the weather will impact the remainder of the operation. Aircraft transit directly to the target or get routed around bad weather. Re-routing will result in less time on station and potentially fewer shots at the target.

Various degradation factors based on weather impact the ability of red and blue forces to detect and engage each other, while random number draws determine the outcome of each of those events. Turns are simulated in five-minute intervals. Once the blue force has destroyed the red force, or failed to engage within the permitted time on station, it returns to the carrier to recycle for another mission. If the aircraft is shot down, it is removed from the simulation. An example of a strike mission and the factors that impact a mission’s success is shown in Figure 7.

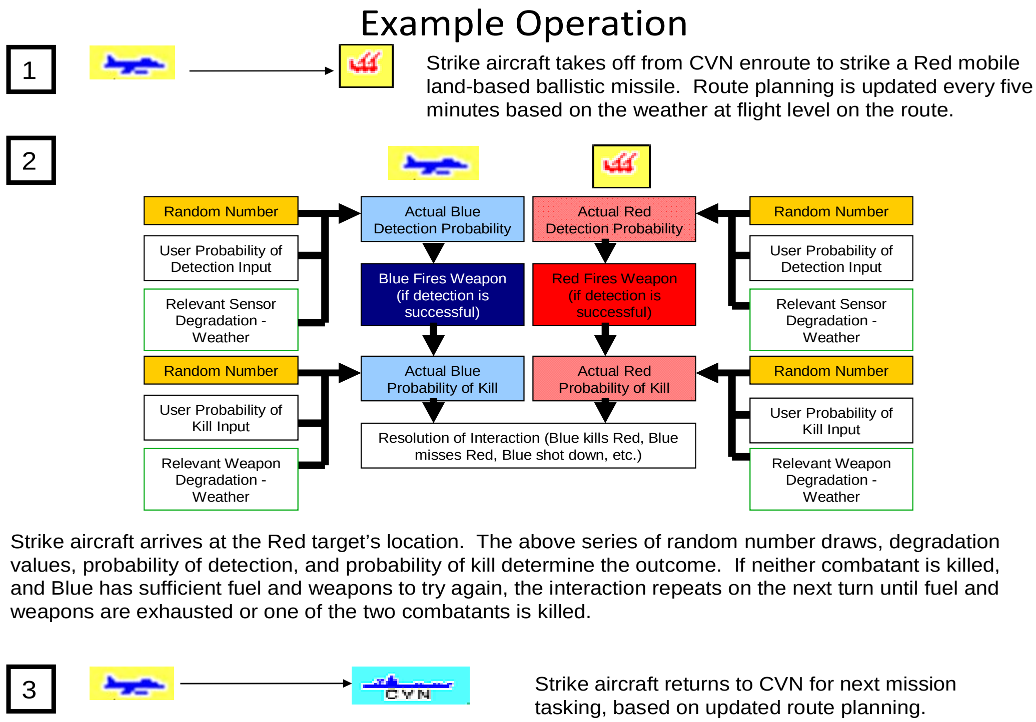

Figure 7. Example of Operational Interaction between Blue and Red Forces.

F. OPTIONS AND FEATURES IN WIAT

WIAT allows the user to choose from among several different methods of simulating forecast error; in this thesis, we use Gaussian errors for weather prediction. WIAT also allows customization of platform/weapon/sensor parameters, environmental degradation factors, METOC thresholds, mission thresholds, and simulation parameters.

WIAT uses METOC threshold parameters of Maximum Yellow and Maximum Green to specify the points at which conditions and forecasts move from one color to another in the red/yellow/green stoplight paradigm; red indicates meteorological conditions that would prevent safe operations, yellow indicates borderline conditions, and green indicates safe conditions. These thresholds are used in calculation of METOC performance metrics, as well as in the determination of sensor and weapon performance degradation.

Mission threshold parameters provide a go/no-go value for the simulated decision maker, and are used determine when changes are made in mission planning in terms of prioritization and/or resource allocation.

Platform, weapon, and sensor parameters allow the user to tailor characteristics of each to mimic realistic capabilities for both Blue and Red forces (i.e., transit and loiter speeds of an F/A-18, number of weapons an aircraft can carry, probability of kill of a given weapon, probability of detection of a given sensor, etc.)

Environmental degradation factors determine the impact that weather conditions above the specified METOC thresholds will have on the effectiveness of a weapon or a detection capability. They work as fractional multipliers of sensor and weapon performance, for both friendly and opposing forces. The weather thresholds that affect various sensors and weapons are shown in Table 1. For example, if visibility, low cloud cover, or mid cloud cover are over their METOC thresholds, blue force UAV sensors will have probability of detection (Pd) multiplied by the environmental degradation factor (a number between 0 and 1). Weapons will suffer reduced probability of kill (P_k) when METOC thresholds are exceeded. Adverse conditions can also lead to increased flight time en route from the CVN to the target, or inability to conduct a battle damage assessment (BDA) once the engagement between Blue and Red forces has concluded.

Weather Parameters and Degradations

		LCD	MCD	HCD	Wind Speed	Visibility	Precipitation
BLUE	UAV Sensors	x	x			x	
	Fighter Sensors	x	x			x	
	Ground Sensors					x	x
	Strike Weapons (Dumb/GPS-guided bombs)				x		
	CAS/KI Weapons (Laser-guided bombs)					x	x
RED	Air Defense Sensors	x				x	
	Ground Sensors					x	x

Table 1. Weather parameters that affect Blue and Red platforms and sensors when actual weather is over METOC thresholds.

Simulation parameters include features such as the number of simulations to run in a block, and the intervals at which weather is displayed in the WIAT graphical user

interface (GUI). One simulation parameter of particular interest to the METOC user is the ability to set a given rate at which mission cancellation recommendations are accepted for both planning and execution forecasts. This feature simulates the level of confidence the decision maker holds in the forecasts provided by METOC personnel. (SPA 2007c)

G. WIAT OUTPUT

As the simulation is run to the conclusion of the campaign period, WIAT collects various MOEs and provides metrics files that detail the performance of Blue forces in prosecuting the campaign. Three primary categories of metrics are compiled for the user: METOC performance metrics, METOC impact metrics, and operational impact metrics. METOC performance metrics (also referred to by the creators as "METOC metrics") compare forecast and observed conditions. Examples are forecast accuracy (FAC), false alarm rate (FAR), and weather event probability of detection (POD, which should not be confused with sensor P_d.)

METOC impact metrics measure the effect that weather forecasts have during mission planning. For example, KI and Strike missions are either changed or not changed based on forecasts. The true weather at mission execution time is compared to the forecasts for that mission to determine whether the forecast-recommended change was correct or not, or if no change was recommended, whether that input was correct or not. If a forecast indicates adverse weather that results in a mission change, and that forecast is verified by the true weather at that later time period, the mission is considered "saved." In contrast, a forecast indicating acceptable weather leading to no mission change, but with adverse weather conditions occurring during the mission, would be considered a mission that "could have been saved." Adverse conditions in only one of the six weather parameters would activate a mission change recommendation, and separate metrics are generated to indicate which of the missions were changed due to adverse conditions for each weather parameter. These metrics are generated for all execution and planning forecasts. (SPA 2007b)

It is important to note that in some METOC impact metrics, the CAS mission is removed from consideration or absorbed into other mission categories. CAS is determined to be the most important mission of the three, in that it involves escort and

support of ground forces in a simulation of a ground campaign in Iraq. This determination yields the philosophy that if CAS is required by ground forces, aircraft will always be sortied to support that mission. Therefore, while there will be other metrics associated with the CAS mission, METOC impacts metrics are effectively eliminated, since forecasters cannot recommend against the mission being flown.

Operational impact metrics focus on the impacts that the actual weather has on the conduct of the campaign. For example, the number of times that adverse conditions led to extended flight times for aircraft to reach a given target is reported as an operational impact metric. This impact can also cause decreased time over the target due to increased flight time en route, inability to determine a firing solution, or inability to conduct a BDA. Metrics are also collected to determine P_k.

H. LIMITATIONS TO WIAT

There are a series of considerations and limitations to the simulated operational decision maker and the campaign that are relevant to understanding how WIAT runs and the metrics it generates. The parameter values themselves have some restrictions. When the user chooses the Gaussian distribution error, WIAT takes the actual weather value at each point and sets it as the mean of the distribution from which forecasts will be produced. The sigma value and a randomly generated number provide the error for that given grid point, which is then applied to the actual value to generate the forecast weather value. However, forecast values can fall outside the realm of reality (i.e., cloud cover less than 0 or greater than 8 octas.) In these cases, WIAT re-assigns the value as the boundary value; a returned forecast of 9 octas would be re-assigned as 8 octas. Additionally, the Gaussian distribution for the cloud cover parameters is discretized—no fractional values are permitted for forecast values (i.e., 1½ octas of cloud cover.) However, fractional values are permitted for wind speed, visibility, and precipitation.

Several aspects of the parameter values also either limit analysis or result in simulation conditions that are less than realistic. Wind speed is treated as one value for the surface, for the (unspecified) altitude en route and over target or the kill box.

Visibility in the ESG database is a function of precipitation and relative humidity. However, dust and aerosols are taken into account in the setting of METOC thresholds. This mismatch is not transparent to the user.

There are several situations in WIAT that do not incorporate realistic campaign planning concerns. First, only the weather at the target is considered when determining whether or not a mission will remain on the ATO and take off as scheduled. The weather en route does impact the length of flight for the aircraft, but cannot cause mission cancellation or the loss of the aircraft; the operational decision maker has no means of factoring an increased flight time into a go/no-go decision. Further, weather conditions at the CVN are not incorporated; in an actual campaign, weather at the launching platform would be an absolutely critical consideration.

Second, while there is a tiered mission assignment process that fills CAS, Strike, and then KI missions, all missions within those subsets are weighted exactly the same in terms of importance. In reality, the war-fighter commonly makes decisions between whether to attack a more important target in adverse weather or a less important target in good weather. This construct is never a possibility in WIAT – adverse weather over a given target will simply cause that particular target to be reassigned to the bottom of the JPITL.

Aircraft also continue to loiter over the target if they cannot sense it, or if they sense and engage the target but do not kill it. There is no mechanism for a mission to change its target during flight; inability to sense the target will force loiter time until the target can be detected, which leads to situations where the Red forces will have multiple opportunities to detect and engage the Blue forces.

A major shortfall of WIAT is that it does not compute metrics for daily results. Therefore, there is no method to determine whether or not the campaign is effectively completed more quickly, or whether or not a given day's events or weather had a proportionately large impact on metrics for the campaign.

I. WIAT IN THIS THESIS

WIAT simulates campaign events that in the real world are subject to random processes, such as whether or not a sensor will detect a target, and whether or not, once it

is fired, a weapon will kill a target. Because WIAT uses random draws to simulate the stochastic nature of these and similar events, two simulations with identical user inputs will not have identical outcomes or identical output metrics. To ensure the metrics produced are representative results of the input conditions, each simulation is run 30 times, and each metric is averaged over the 30 runs.

This study focuses on the comparison between deterministic and stochastic forecasting, with changes in various metrics (instead of ROI) serving as an indication of value. Changes were made to a subset of parameters to reflect each of the deterministic and stochastic scenarios, which will be detailed further in the Methodology chapter. To avoid confusion, this thesis will use the term WIAT* when referring to the author's application of the program, while WIAT will refer to the program in its original context.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. METHODOLOGY

A. WEATHER DURING THE STUDY PERIOD

This study focuses on the six weather parameters previously discussed for 1–7 January 1993, shown in Figures 8–14. The period in question opens with a high-pressure ridge over northern Iran and Turkey, with a relatively weak area of surface low pressure over the eastern Mediterranean Sea. Over the next 24–48 hours, a small surface trough moves southeast along the border between Iraq and Saudi Arabia to the Arabian Gulf. The trough continues east over the Zagros Mountains of Iran, and orographic lifting aids in generating light precipitation over southeastern Iraq and southwestern Iran. The pressure gradient, which weakens with the passage of the low-pressure area on 4–5 January, gradually re-strengthens over 6–7 January as the upper level ridge becomes more pronounced. A surface low also develops over the eastern Mediterranean towards the close of the study period. Visibility is generally poor in northern Iraq throughout the period, with periods of impacts to southern and central Iraq during 3-4 January and again on 6-7 January. Precipitation was a fairly constant feature in the eastern Mediterranean, in the vicinity of the aircraft carrier’s launch area, for the duration of the period with the exception of 5 January. In Iraq, precipitation was extremely sparse and sporadic; the mountains in northern Iraq and the vicinity of Baghdad saw the heaviest precipitation rates, just under 2 mm/hr, associated with the development of the surface low on 7 January. Wind speeds were greatest when the surface ridge-trough pairing was well established and a strong gradient existed between the two features.

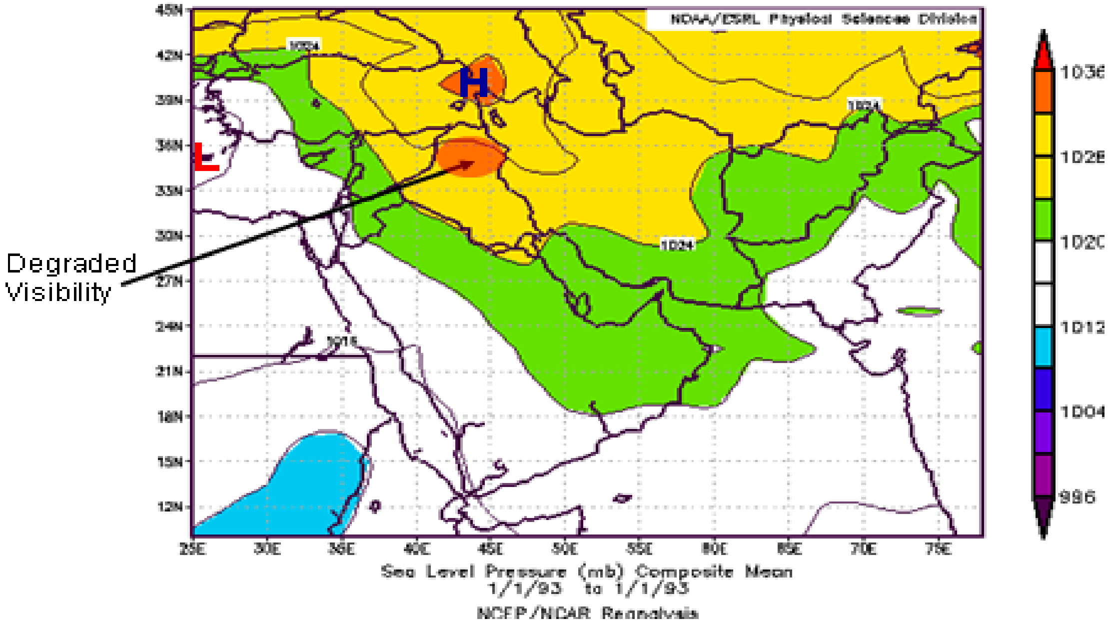

Figure 8. NCEP/NCAR Re-analysis of Sea Level Pressure, 1 January 1993, with noted features of interest.

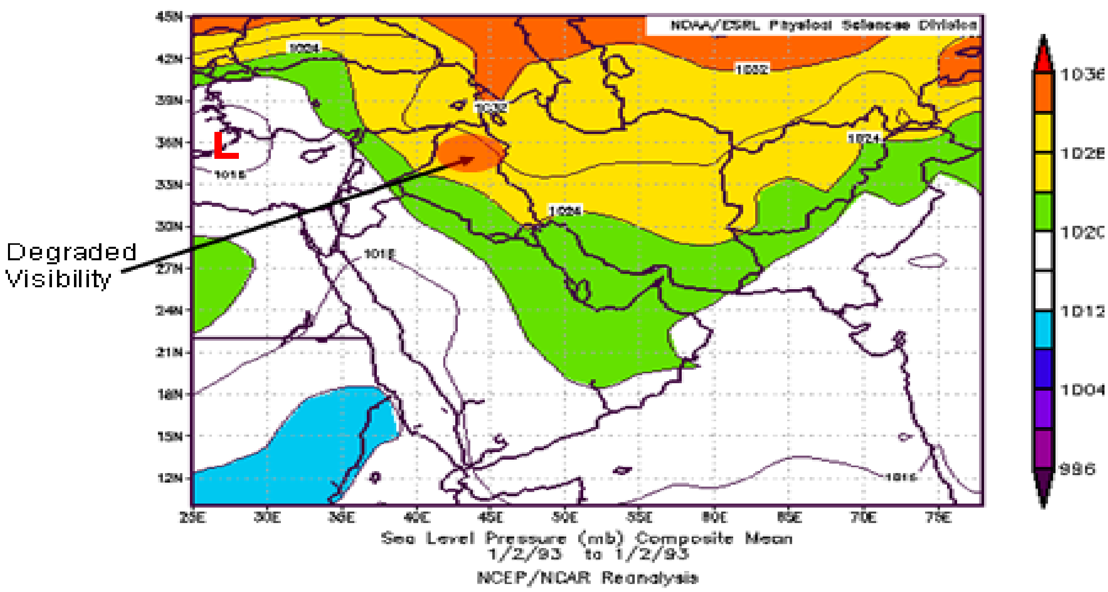

Figure 9. NCEP/NCAR Re-analysis of Sea Level Pressure, 2 January 1993, with noted features of interest.

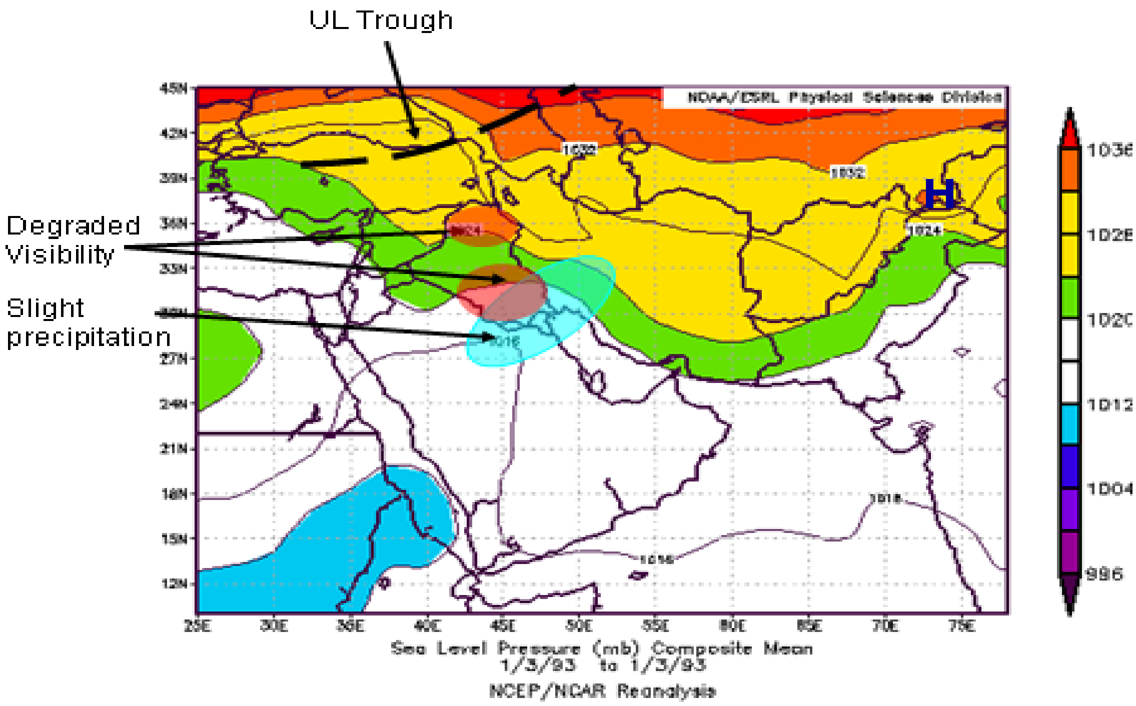

Figure 10. NCEP/NCAR Re-analysis of Sea Level Pressure, 3 January 1993, with noted features of interest.

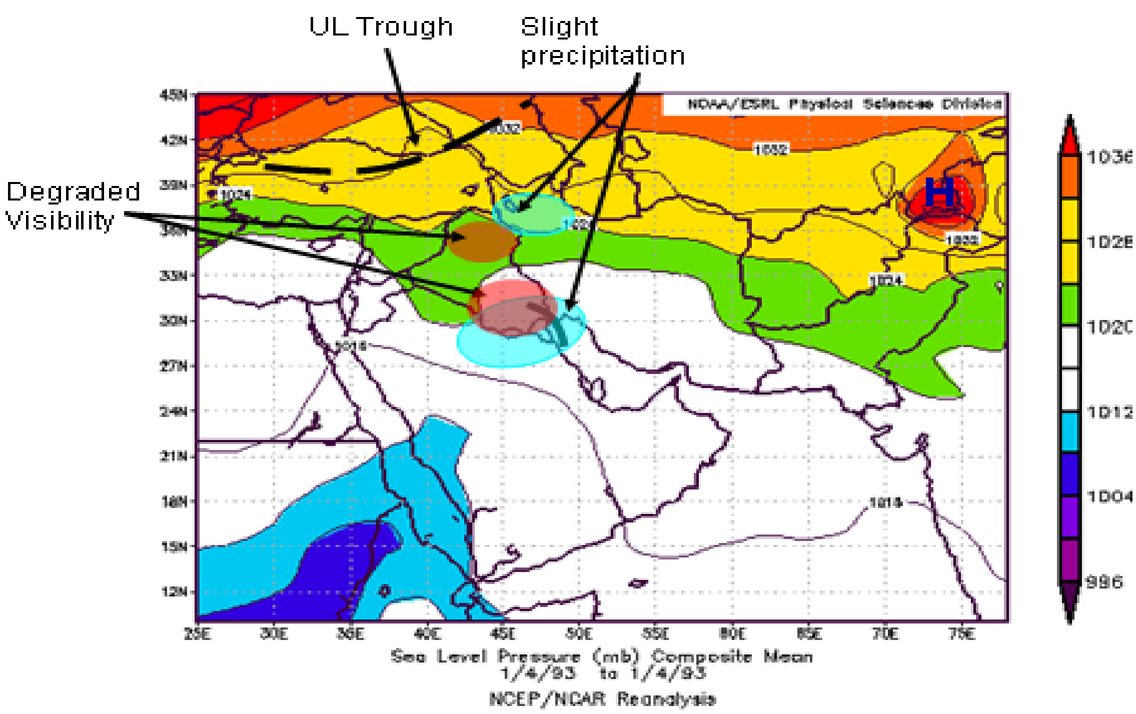

Figure 11. NCEP/NCAR Re-analysis of Sea Level Pressure, 4 January 1993, with noted features of interest.

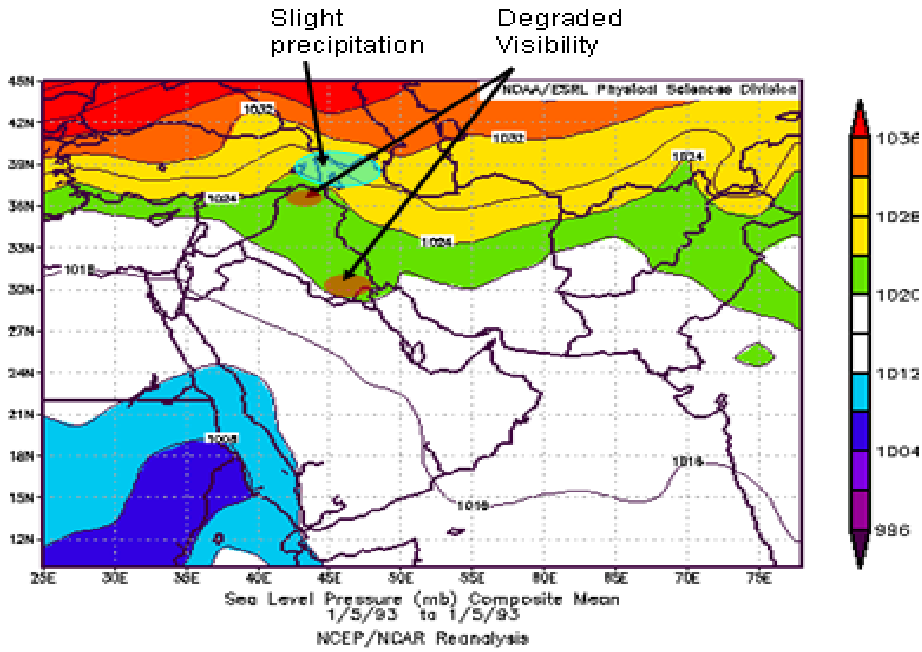

Figure 12. NCEP/NCAR Re-analysis of Sea Level Pressure, 5 January 1993, with noted features of interest.

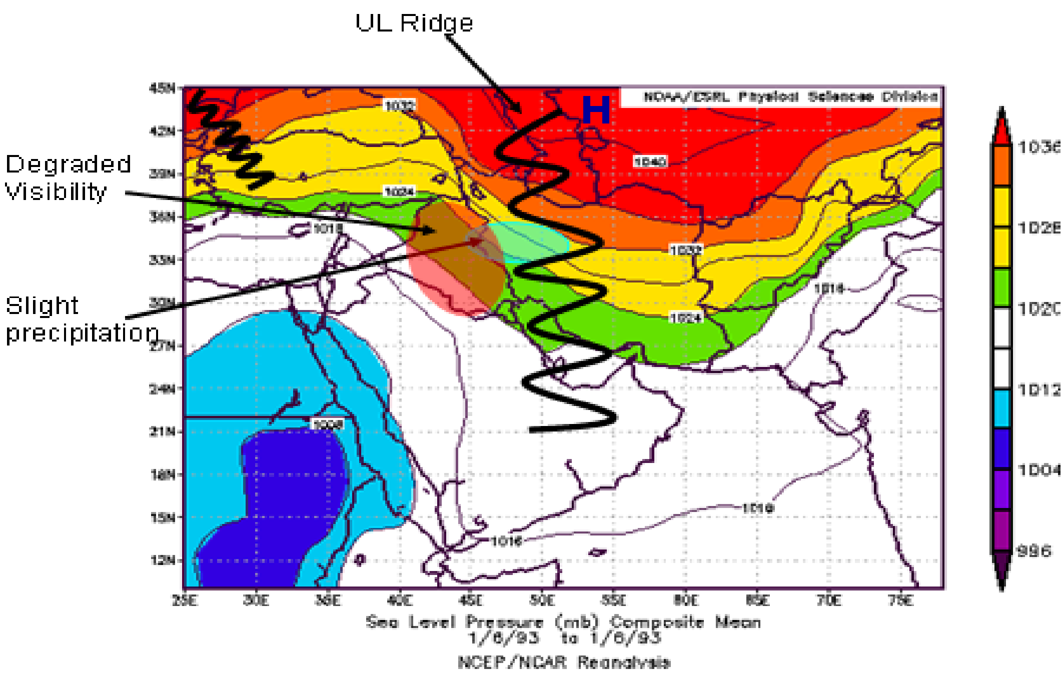

Figure 13. NCEP/NCAR Re-analysis of Sea Level Pressure, 6 January 1993, with noted features of interest.

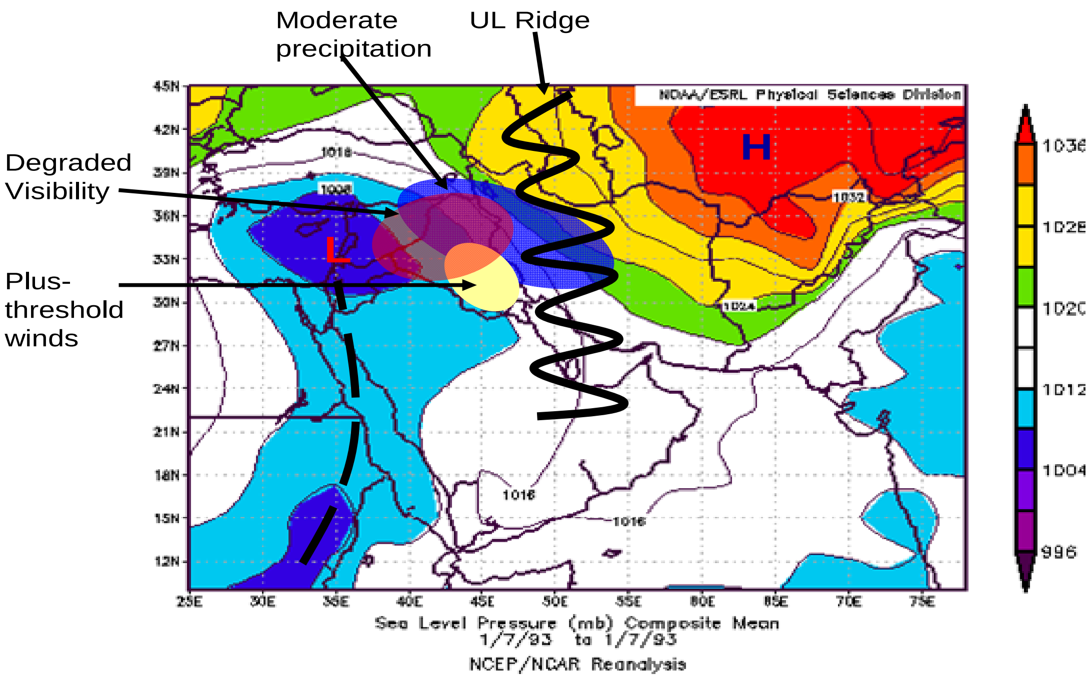

Figure 14. NCEP/NCAR Re-analysis of Sea Level Pressure, 7 January 1993, with noted features of interest.

B. MODIFICATIONS FOR STOCHASTIC WIAT*

In order to isolate the impacts of deterministic and stochastic forecasts, inputs were generally maintained at default WIAT levels. For the deterministic forecasts, the only items that differed from default settings were the rates at which the operational decision maker accepts the METOC recommendations (default levels were 80% for the execution forecast recommendation and 50% for the planning forecast; for this thesis, 100% levels were employed.) The user input settings for WIAT used in the deterministic scenarios are detailed in Appendix B. In our study, mission thresholds of 6 octas cloud cover for all three cloud levels, 11 meters per second wind speed, 9700 meters visibility, and 10 millimeters per hour of precipitation were used in generating the probabilistic forecasts. Decision flow remains the same in WIAT*.

Two very significant changes are made in the inputs for WIAT that enable WIAT* to simulate decisions based on stochastic weather forecasts and probability decision thresholds. First, the weather prediction method shifts from using a Gaussian prediction method to a hand-entered forecast. Instead of applying errors to current

weather to generate forecast values in WIAT itself, the forecasts with appropriate errors are generated in the Ensemble Generation macro; these forecasts are then used to produce the probability file that serves as the weather file for WIAT. Second, the meaning of the KI and Strike mission thresholds are changed from being strictly weather parameters (i.e., 6 octas of cloud cover or 11 m/s wind speed) to probability thresholds. The operational decision maker in WIAT is making the go/no-go decision based on a comparison of the likelihood of the weather at a given location exceeding a METOC threshold with a probabilistic mission threshold. Table 2 shows the probability thresholds, with their equivalent scaled inputs for WIAT*, that were tested by running 30 simulations for each probability threshold.

Probability Thresholds and Corresponding WIAT Mission Threshold Inputs

	LCD	MCD	HCD	Wind Speed	Visibility	Precipitation
100.0%	8	8	8	14.0338	0.00	0.5000
87.5%	7	7	7	12.2796	8738.87	0.4375
75.0%	6	6	6	10.5254	17477.75	0.3750
62.5%	5	5	5	8.7711	26216.62	0.3125
50.0%	4	4	4	7.0169	34955.50	0.2500
37.5%	3	3	3	5.2627	43694.37	0.1875
25.0%	2	2	2	3.5085	52433.24	0.1250
12.5%	1	1	1	1.7542	61172.12	0.0625

Table 2. Probability thresholds and their corresponding scaled values for input into WIAT* simulations.

METOC performance metrics will no longer be useful because WIAT* would attempt to verify each forecast individually. Probabilistic forecast verification must be

against distributions of verifying observations, not individual observations. Alternative metrics, in the form of reliability diagrams, are generated to evaluate the stochastic forecasts; these are presented in Chapter V.

C. SIMULATED ENSEMBLE AND STOCHASTIC FORECASTS

1. Determination of Appropriate Sigma Values

In generating the simulated ensemble forecasts and probability files for WIAT,* we needed to select a set of sigma values. Because our purpose was comparison with WIAT using deterministic forecasts, we needed to choose the same sets of values for both WIAT and WIAT* simulations. Further, we needed to ensure that our sets of sigma values were physically and operationally reasonable. In the deterministic forecasts, sigma represents the standard deviation of forecast error. In the ensemble forecasts, sigma represents the spread of the ensemble, a measure of forecast uncertainty.

In WIAT, the sizes of 1-sigma variations for each of the meteorological parameters are hard-wired (i.e., they cannot be changed with user inputs). For example, 1.0 sigma in WIAT is equivalent to 1 octa of cloud cover and 1.75 meters per second of wind speed. The size of standard forecast errors when using the Gaussian error method was therefore controlled by setting a value for sigma. This value is one of the user inputs; however, there is only one sigma value allowed, so the sizes of standard error for cloud cover and wind speed cannot be changed independently. Table 3 shows meteorological values for each sigma value employed.

Meteorological Meaning				
Sigma Value	LCD/MCD/HCD (Octas)	Wind Speed (m/s)	Visibility (m)	Precipitation (mm/hr)
0.5	0.5	0.875	4300	0.03125
1.0	1.0	1.750	8600	0.06250
2.5	2.5	4.375	21500	0.15625

Table 3. Meteorological meanings for sigma values used based on WIAT assignment of values for 1.0 sigma.

Research that would establish appropriate sigma ranges to use was relatively thin, particularly for precipitation and visibility. We found Sittel's cloud cover verification studies of the Diagnostic Cloud Forecast (DCF) useful, however. CDF forecasts cloud cover over the continental United States from a "statistical relation based on recent performance of [a] mesoscale model" (Sittel, slide 3). Sittel finds that root mean square error (RMSE) values averaged 47% from forecast hours 6 to 48 during June 2009, with little deviation from that average, and no trend line suggesting better forecasting with decreased lead time. Sittel also describes a "20–20 Index," a number between 0 and 1, which for values near 1, indicates a large number of situations where the difference between forecast and observation is less than 20%. The index for June 2009 indicates that there is fewer than 20% difference between forecast/observed pairs 88% of the time, again remaining the same with no appreciable trend from hours 6 to 48.

Morone and Pond (1994) studied the verification for the Rapid Update Cycle (RUC) and their analysis is instructive for establishing appropriate sigma values for wind speed. The RUC model is singular among National Center for Environmental Prediction (NCEP) models in that it is used to provide information on imminent weather conditions; it provides analysis and forecasts out to 12 hours every three hours (Morone and Pond 1994). The verification study found that RMSE for model surface winds, as compared to rawinsonde observations, was between 3 and 4 m/s at the analysis, 3-hour, and 12-hour forecast times.

Based on these studies, as well as limited ability to vary sigma values within WIAT and WIAT*, we chose sigma values of 0.5 and 1.0 for execution forecasts and 1.0 and 2.5 for planning forecasts. The larger sigma value at the planning timeframe incorporates the common notion that the forecaster will generate better forecasts for the execution forecast time frame than for the planning forecast.

2. Generation of Ensemble Files

We created a macro to generate ensemble files to be used by WIAT* in lieu of the deterministic forecast used by WIAT. A macro is a tool within Microsoft Excel that allows direct coding (in the Visual Basic for Excel programming language) of desired functions and references, instead of using standard Excel functions from drop-down

menus and cell references respectively. Direct coding allows Excel to run a series of computations without requiring continued input from the user.

The macro reads actual weather data from the input file and creates forecast data from it using input sigma values. Input sigma values for the low and high uncertainty cases were used to generate the execution and planning forecast values for 30 ensemble members. Corrections were applied to forecasts for values outside forecast boundaries, as described in Chapter II; a separate set of ensemble members that did not use the "wall" represented by the maximum values for wind speed, visibility, and precipitation was also generated but not run through simulations in WIAT*. Graphs in the Results Chapter demonstrate that the forecast distributions were similar, thus eliminating the need to conduct the simulations in WIAT*. The macro also creates a probability file based on how many of the 30 ensemble members exceed parameter thresholds equivalent to the METOC thresholds at each location and forecast lead time. The probabilities in this file are then used as forecasts for WIAT*.

Several alternative means of generating ensemble members were investigated. We tested an ensemble generation method that included applying random errors to the true weather file, in effect treating the ESG values as observed values, and our new values as the true values. We then used the new truth values, applying random errors a second time to generate each ensemble member. Ultimately, we decided against this process, because we wanted the errors of our ensemble members to mimic as closely as possible the errors of the deterministic forecasts produced within WIAT. We also initially modeled a beta distribution for the wind speed, visibility, and precipitation errors, which would have more closely matched the typical distribution of these values. However, we also eventually eliminated the use of the beta distribution to more closely parallel WIAT's forecasting process.

3. Distribution of Weather Values

Weather input for WIAT* in stochastic simulations, as discussed in Chapter IV, consists of scaled probability values for both planning and execution forecasts. Because the forecasts are stochastic, METOC performance and METOC impact metrics are irrelevant for analysis. As a check on the quality of our code, after generating the 30

ensemble members, we compared the distribution of forecast values to the distribution of actual forecast values. Then, we also used reliability diagrams to evaluate how well the forecast probabilities compared with observed rates of occurrence for exceeding the METOC thresholds.

Figures 15–18 show binned distributions of actual weather values and average counts of forecast values for the two forecast periods, in both low and high uncertainty scenarios. As is expected, the distribution of execution forecasts is generally closer than the planning forecasts to the actual weather. Both forecast distributions demonstrate a smoothing of peaks, as compared to the distribution of actual values, and the smoothing increases as more error is introduced. Figures 15 and 16 show that low cloud cover has peak values at the boundary values of 0 and 8 octas, with a smaller peak at 4 octas. Higher levels of uncertainty at increasing lead times yield higher accumulations at non-peak values.

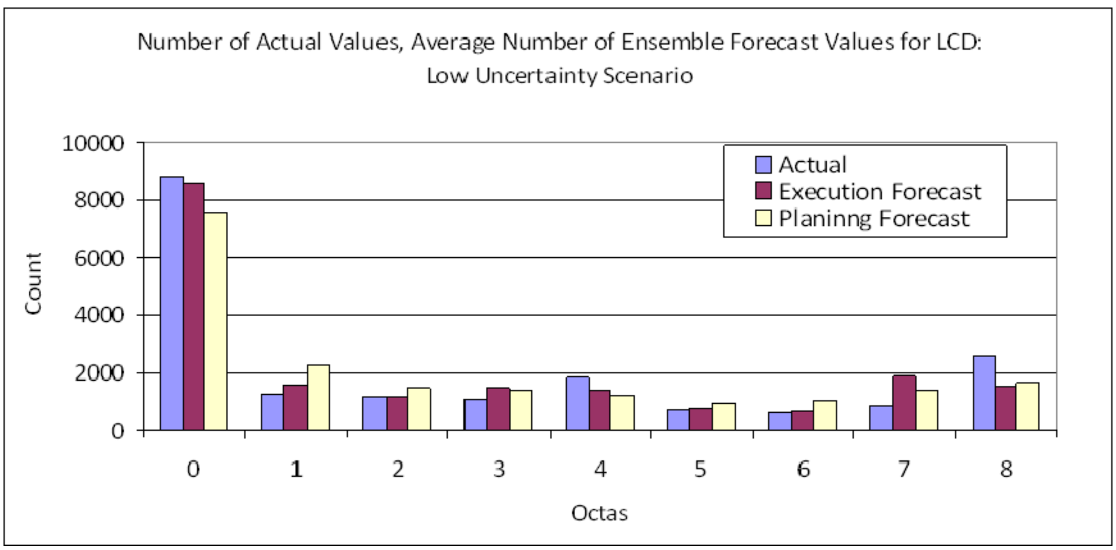

Figure 15. Number of Actual Values, Average Number of Ensemble Forecast Values for LCD: Low Uncertainty Scenario.

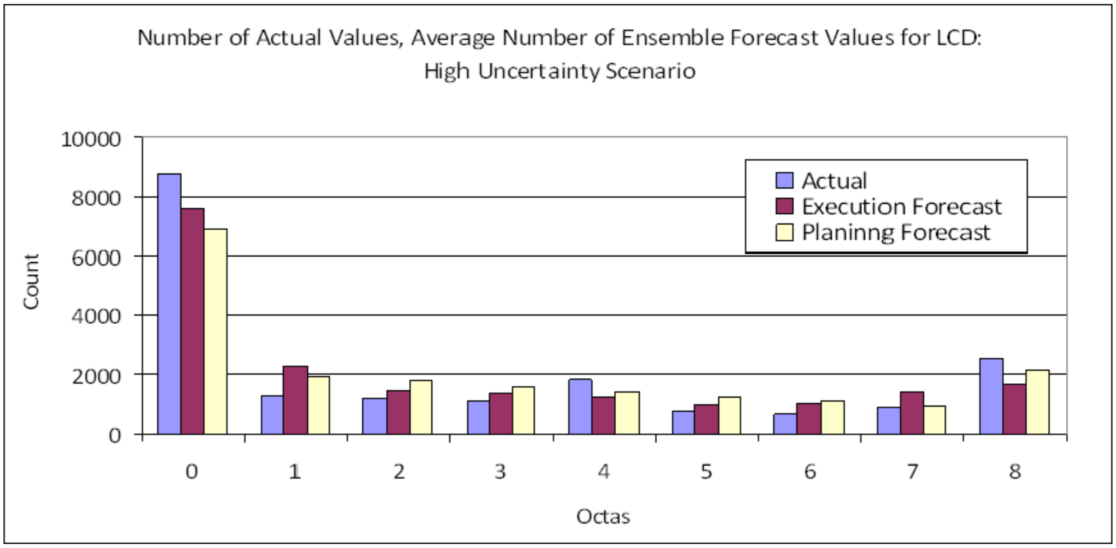

Figure 16. Number of Actual Values, Average Number of Ensemble Forecast Values for LCD: High Uncertainty Scenario.

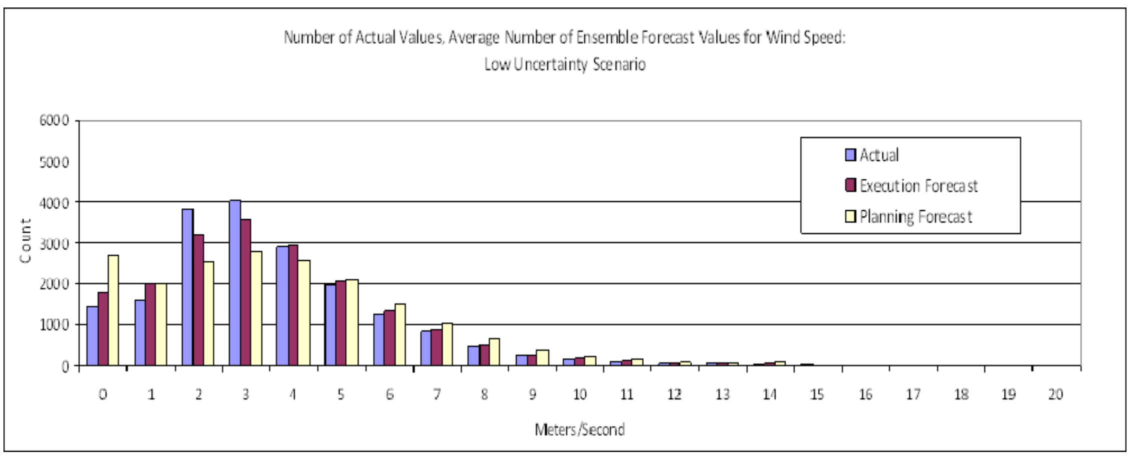

Figure 17. Number of Actual Values, Average Number of Ensemble Forecast Values for Wind Speed: Low Uncertainty Scenario.

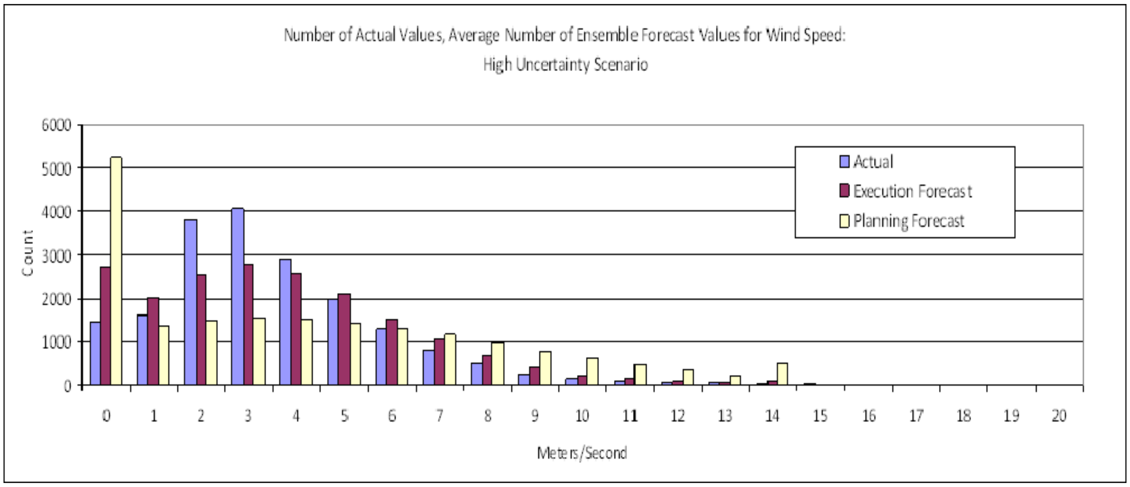

Figure 18. Number of Actual Values, Average Number of Ensemble Forecast Values for Wind Speed: High Uncertainty Scenario.

Figures 17–18 show similar diagrams for wind speed. In these cases, peak values occur at 3-4 m/s, and more spreading of the distribution occurs as more forecast uncertainty is introduced. Peaks in execution and planning forecasts also occur at the boundary value of zero. Because WIAT adjusts returned forecast values that are lower than 0 m/s to a value of 0 m/s, the ensemble generation program used in this study follows the same procedure. As is expected, however, diminishing uncertainty means fewer values returned less than zero, and hence lower accumulation of values at the boundary value.

An additional consideration was taken in creating the scaled probability file that serves as the base weather file for stochastic runs in WIAT*. WIAT will accept fractional octa values for actual weather values, but will not do so for forecast values: it instead truncates the value to an integer. For example, if the actual weather value for a grid location is 6.2 octas, and applying the Gaussian-distributed error yields a forecast value of 6.8 octas, WIAT will truncate the 6.8 octas value to 6 octas, artificially re-distributing the value below the actual weather value. Further, WIAT interprets mission threshold values as the maximum permissible value. In this manner, all forecast values

between 6 and 6.99… octas would be returned to a value of 6 octas, and forecast values intended to be over the threshold would be interpreted as equal to the threshold, and would not result in a mission change.

To remedy this, an additional adjustment to forecast values from the probability file for each scenario was performed that effectively rounds the forecast value up to the next integer. The distributions of values after this adjustment are shown in Figures 19–20. Actual values were not adjusted, but the chart demonstrates how filtering true values in the same manner would yield a different distribution. Concentrations of 0 and 1 would vary significantly between actual and forecast files, but because the values of concern are those from 6 to 8 octas, this variation is acceptable. As it stands, forecast distributions above the threshold value parallel the actual values very closely post-adjustment.

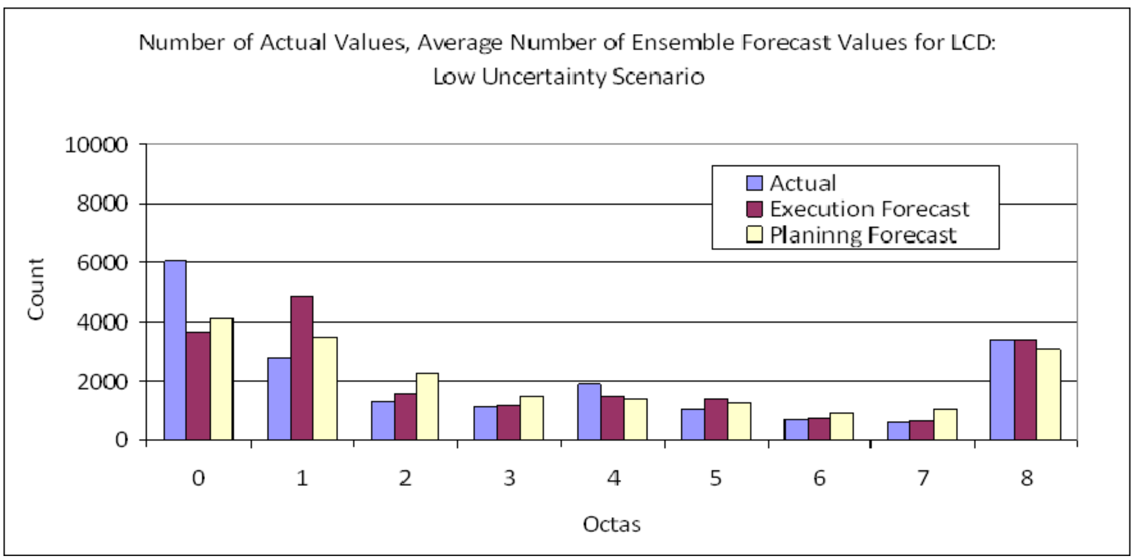

Figure 19. Number of Actual Values, Average Number of Ensemble Forecast Values for LCD: Low Uncertainty Scenario, after adjustment to counter WIAT cloud cover value truncation.

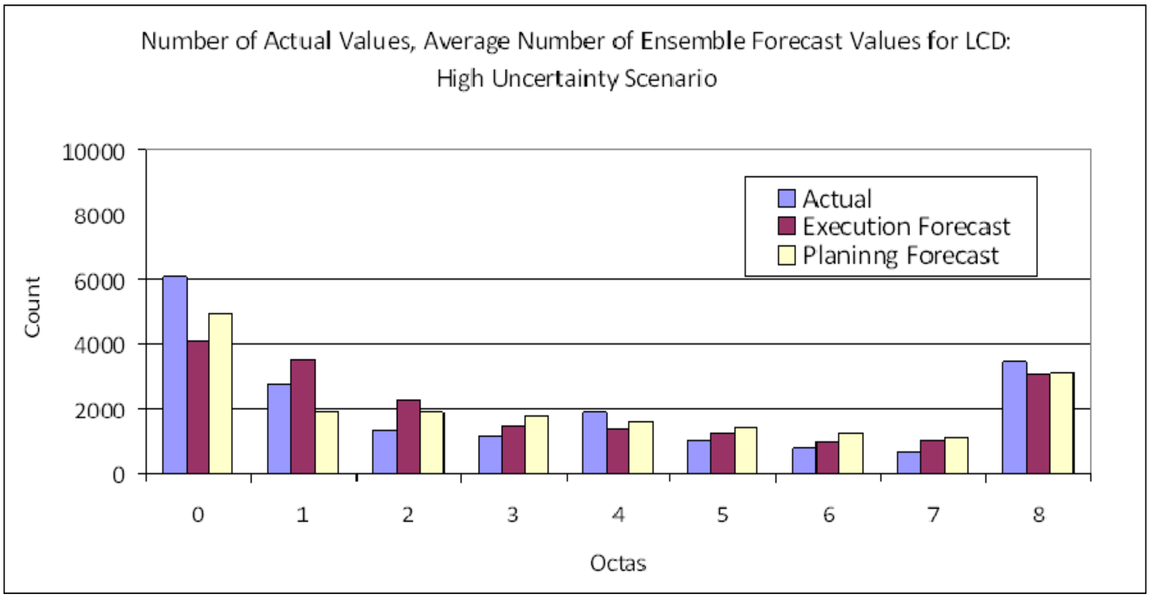

Figure 20. Number of Actual Values, Average Number of Ensemble Forecast Values for LCD: High Uncertainty Scenario, after adjustment to counter WIAT cloud cover value truncation.

V. COMPARISON OF RESULTS FROM WIAT AND WIAT*

A. DETERMINISTIC WIAT RESULTS

1. METOC Impact Metrics

The simulation blocks for the low and high error scenarios yielded METOC impact metrics that show benefit from higher accuracy. Two examples of the METOC impact metrics are shown in Table 4. As either planning or execution sigma values increase, the number of missions incorrectly changed and incorrectly not changed increases. We would expect that higher degrees of forecast error would lead forecasters to unnecessarily change the mission when no change was required (also known as a "false alarm") or allow the mission to proceed when it should have been changed (also known as a "miss").

Selected METOC Impact Metrics from Deterministic Scenarios

	Low Error	High Error
Average Number of Missions Incorrectly Changed	9.97	19.13
Average Number of Missions Incorrectly Not Changed	18.83	32.50

Table 4. Average number of missions incorrectly changed and incorrectly not changed for low and high error scenarios in WIAT.

2. Operational Impact Metrics

The operational impact metrics shown in Table 5 do not show as clear a trend as is evident in the METOC impact metrics. While STK and KI/CAS kills per 100 missions increase when forecasts incorporate less error, as should be expected, the difference between the two is only on the order of 1–2%. This difference is statistically significant at a 95% confidence interval for STK kills per 100 missions (based on one-tailed t-tests.) However, the difference in KI/CAS kills per 100 missions is not statistically significant at that confidence interval.

Selected Operational Metrics from Deterministic Scenarios

	Low Error	High Error
STK Kills per 100 Missions	62.73	61.62
KI/CAS Kills per 100 Missions	26.41	25.76

Table 5. Strike and KI/CAS Kills per 100 missions for low and high error scenarios in WIAT.

Most operational metrics did not demonstrate identifiable trends for analysis. For many of the operational impact metrics, the forecast weather does not actually have a direct impact on the fields used to generate the metrics; often, the actual weather is what causes an impact to Blue forces trying to complete missions. For example, the weapons probability of kill metric is calculated directly from the number of targets killed and the number of weapons deployed against those targets. The number of weapons deployed is derived from whether or not the Blue forces have detected and then destroyed the target, both of which are influenced by actual weather and not a forecast. Assuming that the degradation factor due to weather influences affects Blue and Red to the same degree, even metrics such as the number of Blue manned aircraft shot down cannot be attributed to good or bad forecasts, but instead to the actual weather and the capabilities given to Blue and Red by the user.

B. STOCHASTIC WIAT* RESULTS

1. Reliability Diagrams and Calibration

One of the constant concerns for the operational forecaster is determining whether or not the forecasts issued are good portrayals of the environment. For stochastic forecasts, the forecaster is concerned with reliability and resolution. In order to have stochastic forecasts that are considered reliable, the forecast probability of a given event should equal the observed relative frequency of the event. Resolution refers to how well the probabilistic forecast distinguishes between events and non-events.

Reliability diagrams compare forecast probability with observed relative frequency. For the stochastic scenarios in WIAT*, a macro was generated that counted the number of times a given probability of exceeding the mission threshold was issued for execution and planning forecasts, and compared it to the count of occasions for corresponding valid times and locations where the actual weather exceeded the mission threshold. We would expect a perfectly reliable forecast to follow the straight line shown in the following charts. The introduction of random forecast error, as well as some of the constraints applied in the formulation of the ensemble members, will yield forecasts that are less than perfectly reliable.

Figures 21–22 show examples of reliability diagrams for the low and high uncertainty scenarios; remaining diagrams can be found in Appendix D. Both demonstrate a series of forecast-observed rate pairs that reflect an over-statement of the likelihood of unlikely events, and an under-statement of the likelihood of likely events. For example, Figure 21 depicts that when forecasts indicate a 30% or lower likelihood of exceeding the mission threshold, the corresponding actual weather value never exceeds the threshold; meanwhile, for forecast probabilities 66.67% and higher, we see that the threshold is exceeded in every case. Very small numbers of forecast occurrences away from the extreme probabilities of 0% and 100% accentuate the divergence of the ensemble from forecasts with perfect reliability. Error bars show two standard deviations above and below the observed rate of occurrence.

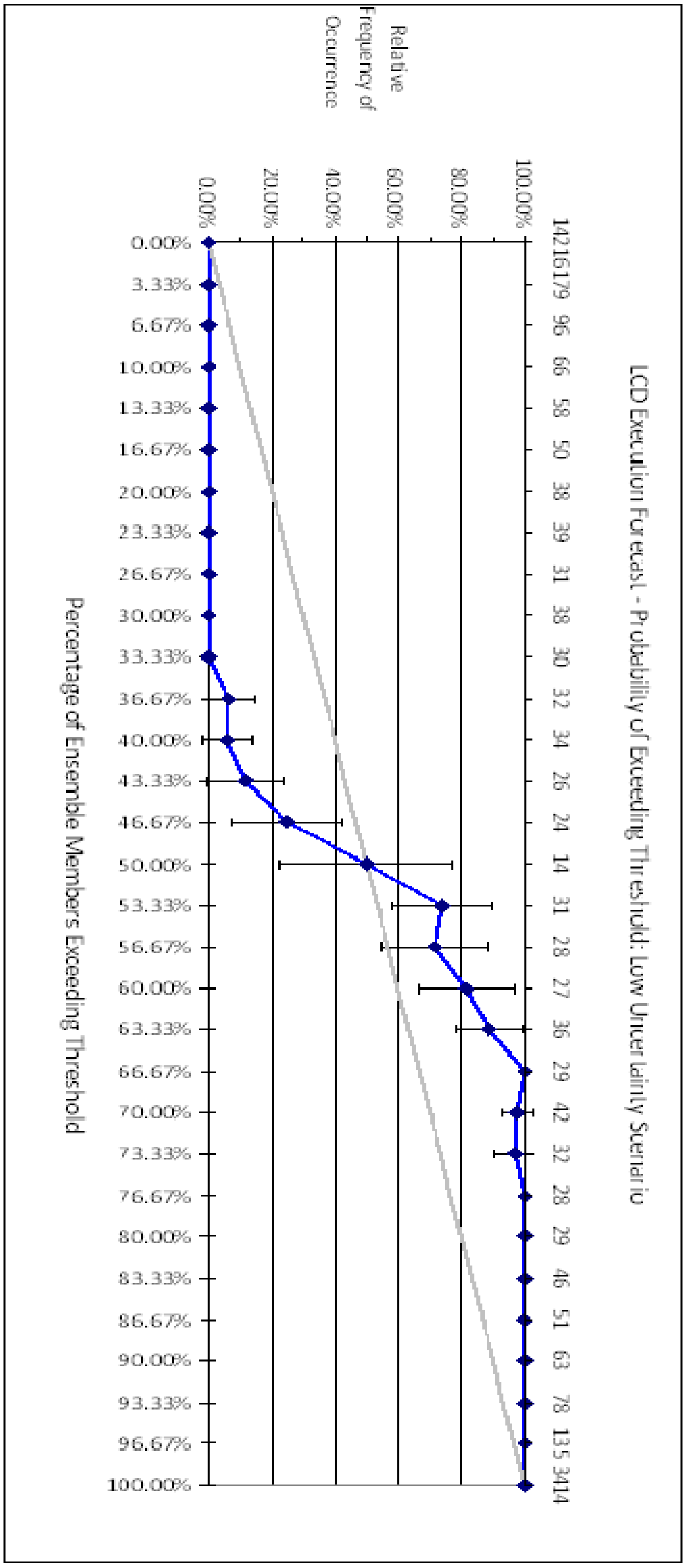

Figure 21. Reliability Diagram with Error Bars for LCD Execution Forecast, Low Uncertainty Scenario.

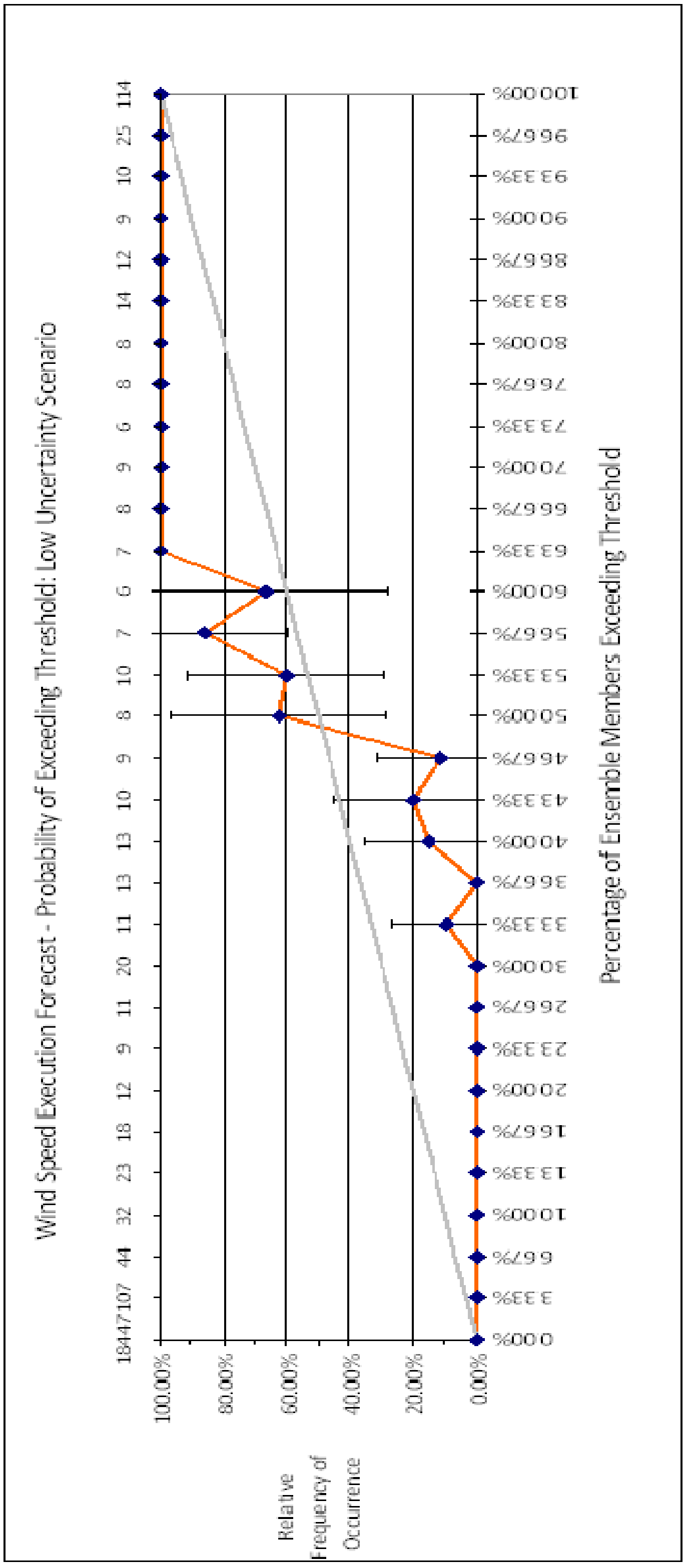

Figure 22. Reliability Diagram with Error Bars for Wind Speed Execution Forecast, Low Uncertainty Scenario.

As seen in Figures 21–22, our ensemble generation technique did not produce reliable probability forecasts. We made a few attempts to change aspects of the ensemble generation macro that we thought might be responsible for the ensembles poor calibration, but were unable to improve the ensemble's reliability. In order to correct the forecast probabilities of exceedance, we employed a calibration process as discussed in Chapter II. We replaced the likelihood derived from the set of 30 ensemble members with the environmental rate of occurrence associated with that probability. WIAT* runs for each threshold level used for the uncalibrated forecast were conducted for the calibrated forecasts as well.

2. Operational Impact Metrics

As in the deterministic WIAT runs, stochastic WIAT* metrics are relevant for analysis; unfortunately, there are few operational metrics that show definitive differences among the various threshold levels, and many other metrics are difficult to explain. Tables 6–9 show several operational metrics from WIAT* runs at each of the eight mission threshold levels, for low and high uncertainty, and for uncalibrated and calibrated probability forecasts. Additional operational impact metrics are tabulated in Appendix F.

Comparing the metrics for the various mission thresholds used, we cannot identify an optimum decision threshold. Further, while it seems reasonable that using a higher threshold (i.e., being willing to fly missions even when there is a very high chance of mission degradation or failure due to weather), could explain the lower number of strike kills per 100 missions, the opposite trend is observed in the KI/CAS kills per 100 missions.

Comparisons between the low and high uncertainty scenarios generally show better strike kill rates, fewer missions impacted, and fewer mission impacts at threshold levels of 87.5%, 75%, and 62.5%, but worse rates for the lowest three thresholds. KI/CAS kills per 100 missions demonstrates the opposite trend in this comparison as well.

Operational Impact Metrics for Stochastic Low Uncertainty Scenario (Uncalibrated)

Metric	Threshold							
	100.0%	87.5%	75.0%	62.5%	50.0%	37.5%	25.0%	12.5%
Strike Kills per 100 Missions	65.55	63.23	63.11	63.35	64.08	65.66	68.42	68.14
KI / CAS Kills per 100 Missions	28.04	25.35	25.65	25.33	25.71	24.77	24.05	22.80
Missions Impacted	165.10	155.07	154.30	151.50	148.73	142.20	127.50	120.20
Total Mission Impacts	205.40	179.17	177.57	176.37	171.10	167.57	152.80	144.37

Table 6. Operational impact metrics for WIAT* runs. Metrics are for the low uncertainty scenario, uncalibrated probability forecasts.

Operational Impact Metrics for Stochastic High Uncertainty Scenario (Uncalibrated)

Metric	Threshold							
	100.0%	87.5%	75.0%	62.5%	50.0%	37.5%	25.0%	12.5%
Strike Kills per 100 Missions	65.55	62.12	61.73	62.94	64.45	67.07	68.80	69.28
KI / CAS Kills per 100 Missions	28.04	27.66	26.92	25.39	25.62	25.03	23.54	21.61
Missions Impacted	165.10	158.67	156.87	152.97	147.63	131.90	120.33	109.13
Total Mission Impacts	205.40	190.20	183.93	178.50	172.20	156.60	144.70	129.27

Table 7. Operational impact metrics for WIAT* runs. Metrics are for the high uncertainty scenario, uncalibrated probability forecasts.

Operational Impact Metrics for Stochastic Low Uncertainty Scenario (Calibrated)

Metric	Threshold							
	100.0%	87.5%	75.0%	62.5%	50.0%	37.5%	25.0%	12.5%
Strike Kills per 100 Missions	65.55	63.16	63.40	63.52	64.18	65.20	66.34	67.04
KI / CAS Kills per 100 Missions	28.04	26.49	26.59	26.69	25.27	24.51	25.43	24.39
Missions Impacted	165.10	155.67	154.40	152.50	148.43	143.40	138.50	130.63
Total Mission Impacts	205.40	178.40	176.37	177.10	171.90	166.03	160.37	151.90

Table 8. Operational impact metrics for WIAT* runs. Metrics are for the low uncertainty scenario, calibrated probability forecasts.

Operational Impact Metrics for Stochastic High Uncertainty Scenario (Calibrated)

Metric	Threshold							
	100.0%	87.5%	75.0%	62.5%	50.0%	37.5%	25.0%	12.5%
Strike Kills per 100 Missions	65.55	62.99	63.14	63.42	64.14	65.61	66.99	68.32
KI / CAS Kills per 100 Missions	28.04	26.09	25.88	25.27	25.97	25.03	24.77	24.25
Missions Impacted	165.10	156.07	152.63	149.93	147.93	142.07	134.13	126.77
Total Mission Impacts	205.40	182.13	177.23	176.23	172.13	166.87	155.90	151.07

Table 9. Operational impact metrics for WIAT* runs. Metrics are for the high uncertainty scenario, calibrated probability forecasts.

Comparing metrics for simulations that used calibrated forecasts to the metrics for simulations that used uncalibrated forecasts, none of the metrics, the number of kills per 100 missions, missions impacted, or total mission impacts, showed a clear benefit to use of our calibration technique. No region of optimum mission threshold appeared with calibration.

VI. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

Results across the spectrum of stochastic WIAT* runs may show improved metrics for stochastic decision-making and weather support. For example, the number of strike kills per 100 missions for the high error deterministic WIAT runs was 61.62 (Table 5) while the strike kills per 100 missions for the high uncertainty stochastic WIAT* runs ranges from 65.55 to 69.28 (Tables 7 and 9). This is an increase of approximately 4-6% over the range of thresholds tested. While this percentage may not be a dramatic improvement, each additional success per 100 missions can bring earlier air dominance/supremacy, shorter campaign lengths, fewer expenses incurred for operation of both man and machinery, and reduced chances of friendly forces being shot down. Our somewhat mixed results do not allow us to definitively conclude that stochastic forecasting provides benefit to the simulated operational decision maker in this campaign scenario, but we cannot rule it out, either. Several primary issues in WIAT prevent us from forming more certain conclusions: the construct of the program itself, aspects of the simulation, and a lack of full transparency as to how metrics are accumulated. However, the limited improvement we did see suggests that further study should be pursued, by implementing different strategies for testing purposes.

WIAT was designed primarily to allow users to test the impacts of varying degrees of deterministic forecast quality in the context of a campaign. Restrictions and adjustments of data in some parts of the program make perfect sense in light of this design aim; however, several features—with the requirement for cloud cover to be in integer octa values the prime example—make adapting the program for stochastic purposes extremely difficult. The choice to truncate forecast values, while not performing the same function on actual values, adds processing steps to the effort. Finally, hard-coding of certain sigma values and maximum parameter values into WIAT limits the range of scenarios where WIAT would be useful.

While it would be unrealistic to expect the programmers of WIAT to have generated a simulator on par with video games such as the *Harpoon* series, there are multiple instances of modeling in WIAT that are less than realistic. In order to parallel the experience, most afloat METOC division officers have to deal with, it seems reasonable to expect that weather conditions at the CVN, and not just the target location, should be used in determining whether a mission will go as scheduled. Weather en route to the target is also a complexity that the decision maker cannot integrate into the planning cycle. There is no capability to prioritize missions by importance, and hence the uniform weighting by decision makers of how to allocate their resources (i.e., to target the less important Red force under good weather, or the more important Red force under bad weather) and the different-valued outcomes of these missions are not factors in WIAT's JPITL and metrics. Finally, while metrics are broken down by execution and planning forecast time frame, the success of missions is not broken down by day. There is no way to tell in WIAT whether or not the majority of Red force kills, for example, were achieved in Days 1 through 3 of the campaign, or evenly spread out over the seven days. This would be a useful statistic, given the fact that there are more missions over the seven-day campaign than there are opportunities to fly them, and that the user is unable to realize a shorter campaign with more accurate forecasts (or even perfect detection and weapons employment.) Further, some metrics seem to only capture impacts of the weather itself, rather than the forecaster's contribution to the decision maker.

Because WIAT employs GCAM-CTS as part of its package, the programming behind the simulation is often a bit obscured. Were it not for the willingness of one of the original programmers to donate considerable time and effort, much would have remained misunderstood. Some elements of WIAT's programming are more accessible, but if the user wants to understand the underlying how's and why's of WIAT architecture, it requires extensive study of programming in Visual Basic for Applications (VBA) in Microsoft Excel. The fact that VBA does not appear to be widely used in the scientific community also limits the avenues for information technology (IT) support.

This particular weather period may not have been a good choice for testing the benefits of stochastic forecasting. As discussed in Chapter IV, weather conditions during

this period were fairly mild; three of the six parameters studied were effectively non-factors. While cloud cover was present for most of the time frame at the low- and mid-levels, the most challenging conditions for a strike campaign would have revolved around the consistently poor weather in the vicinity of the CVN, which of course was not considered in launch decisions. A more challenging weather scenario, with extensive thunderstorms and heavy rain or with shamal wind conditions may have exercised the possibilities more thoroughly.

B. RECOMMENDATIONS

This thesis prompts several recommendations for future study. WIAT could provide an excellent means for further study with several changes to structure and output. First, WIAT should be adapted to incorporate stochastic forecasting with its own set of inputs and metrics. Reliability diagrams would be an appropriate METOC performance metric for use with the stochastic forecast method. Adaptation would also eliminate the need for scaling probability values, as well as permitting finer sampling of probability thresholds in search of optimal values. Incorporating more realistic penalties and complicated degradation schema would also enhance the suitability of the campaign simulation. While the user ultimately can search out the scenario of choice, inclusion of more challenging weather data files might prove more useful to the user than the current set. Users should also have the capability to choose “simpler” forecast simulations that may only consider one weather parameter, run a shorter campaign length, or study only one mission type. Lastly, allowing users to stipulate different sigma values for different weather parameters would enhance the realism of the error distribution. Training on the program, which has been offered by SPA in the past, provides an extremely valuable introduction to the program.

Future study should also attempt to find realistic ways to simulate the generation of ensemble data. Perturbing the actual parameter values to generate the true weather PDF discussed in Chapter II was pursued earlier in the process of this thesis, but ultimately was removed because this constituted adding a source of error where WIAT had no comparable error introduction. Results from those forecasts did not yield more

reliable forecasts than those described in this thesis, but a more correct ensemble generation process should incorporate the notion that observed weather itself contains error. The computing power that would be required to support these sorts of investigations mandates the use of a more flexible programming language; while Microsoft Excel and its version of VBA have proven themselves useful in many contexts, a program such as MATrix LABoratory (MATLAB) would likely prove far easier to use and more capable.

The most exciting research may lie in adapting simulations in WIAT's image to other warfare areas. The large search areas characteristic of ASW, and the relatively micro-scale concerns typical of NSW, would each be suitable and of interest for weather impacts simulations. Further, either of these warfare areas would expand stochastic forecasting studies to oceanographic applications and campaigns; expanding work in stochastic forecasting to oceanographic applications can only serve to enhance the Navy's capabilities in the years ahead.

APPENDIX A: STUDY PERIOD WEATHER

A. SEA LEVEL PRESSURE CHARTS, 1–7 JANUARY 1993

This section shows NCEP/NCAR re-analyses of sea level pressure for the Middle East region in general, and specifically the operating area in the campaign scenario in WIAT.

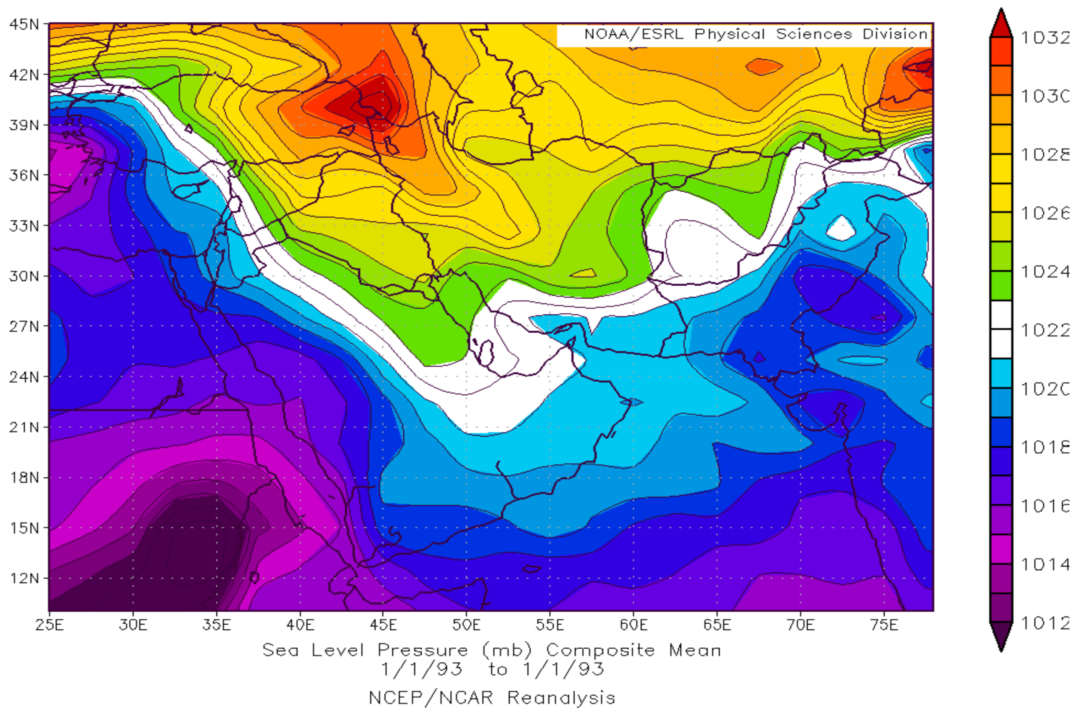

Figure 23. Sea Level Pressure, 1 January 1993.

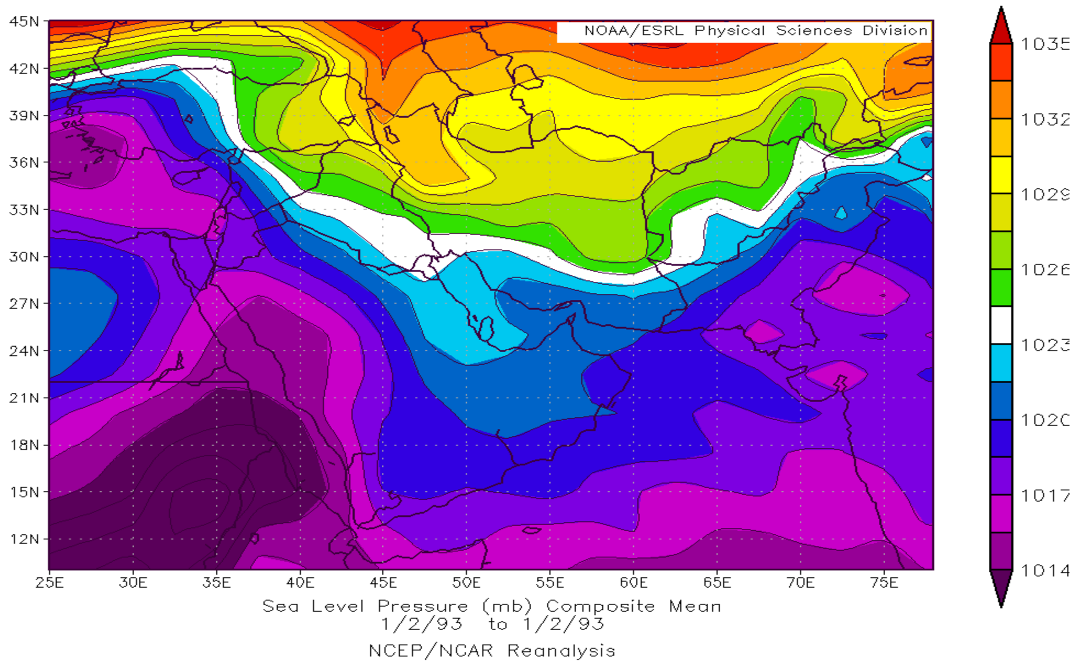

Figure 24. Sea Level Pressure, 2 January 1993.

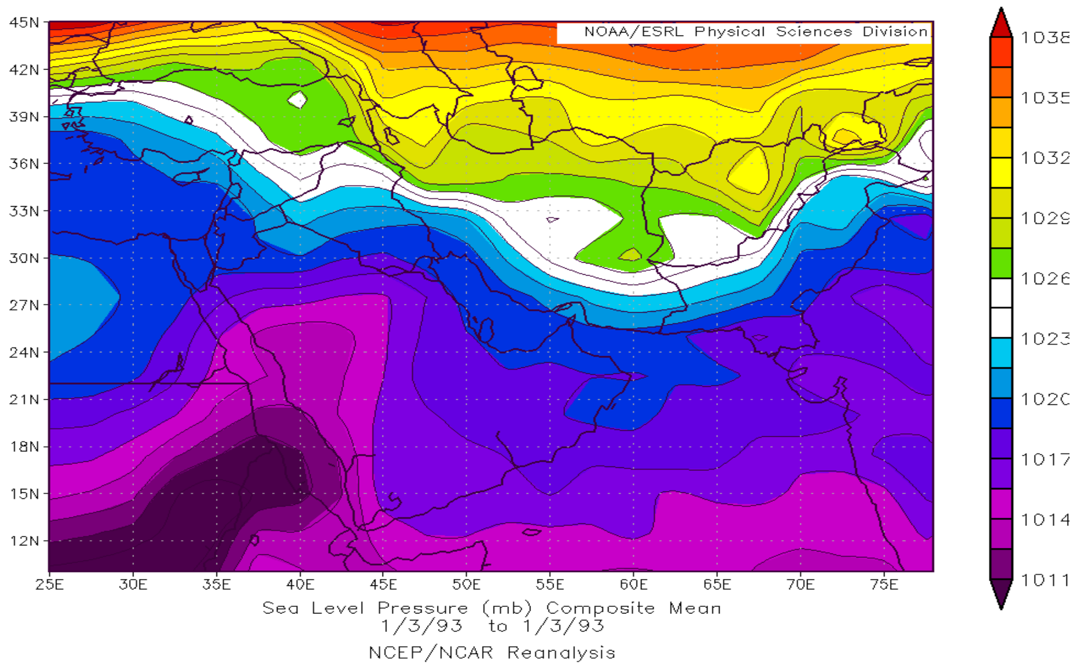

Figure 25. Sea Level Pressure, 3 January 1993.

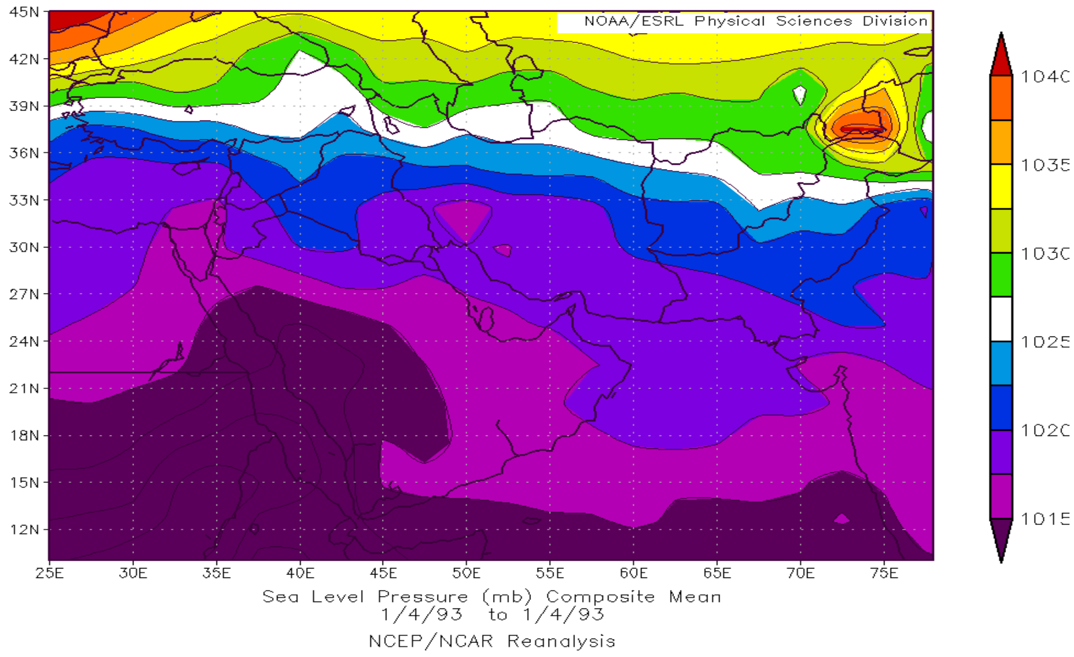

Figure 26. Sea Level Pressure, 4 January 1993.

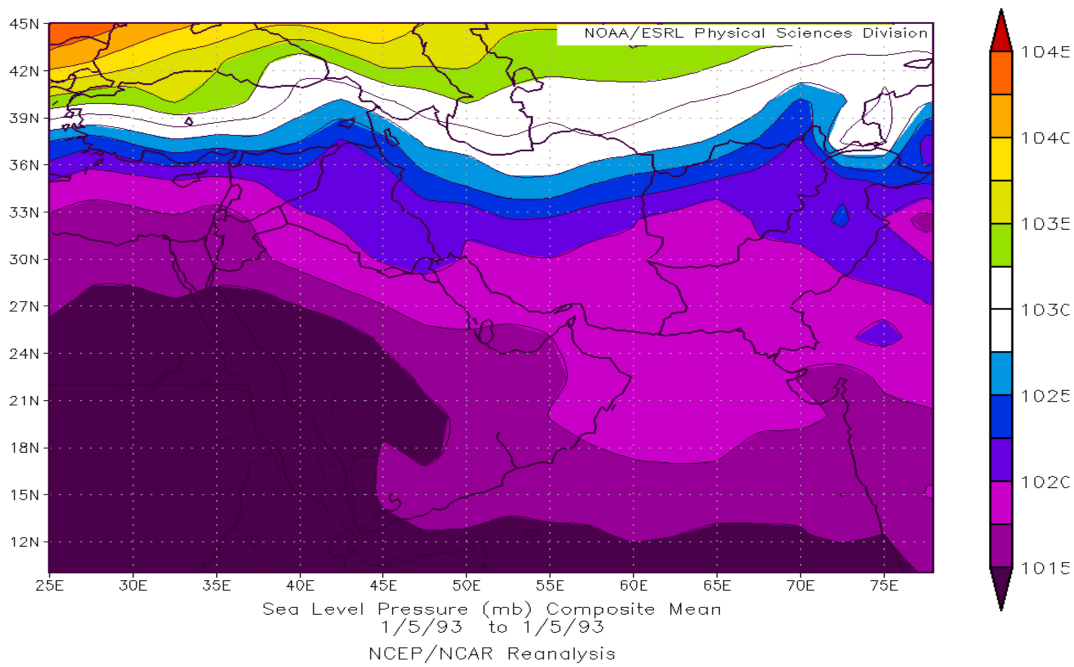

Figure 27. Sea Level Pressure, 5 January 1993.

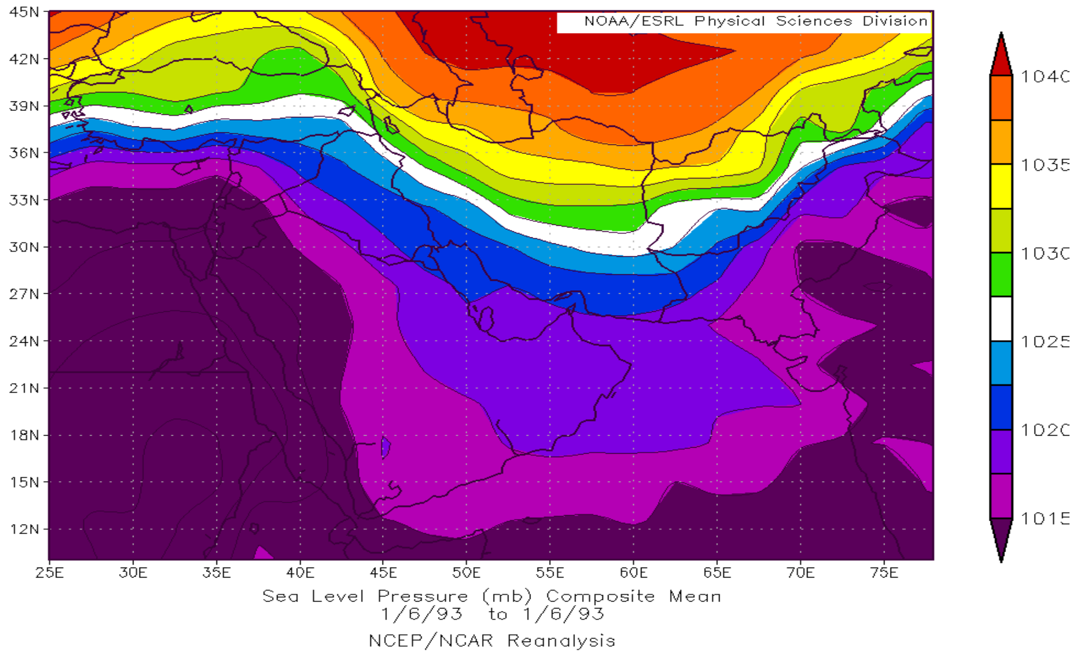

Figure 28. Sea Level Pressure, 6 January 1993.

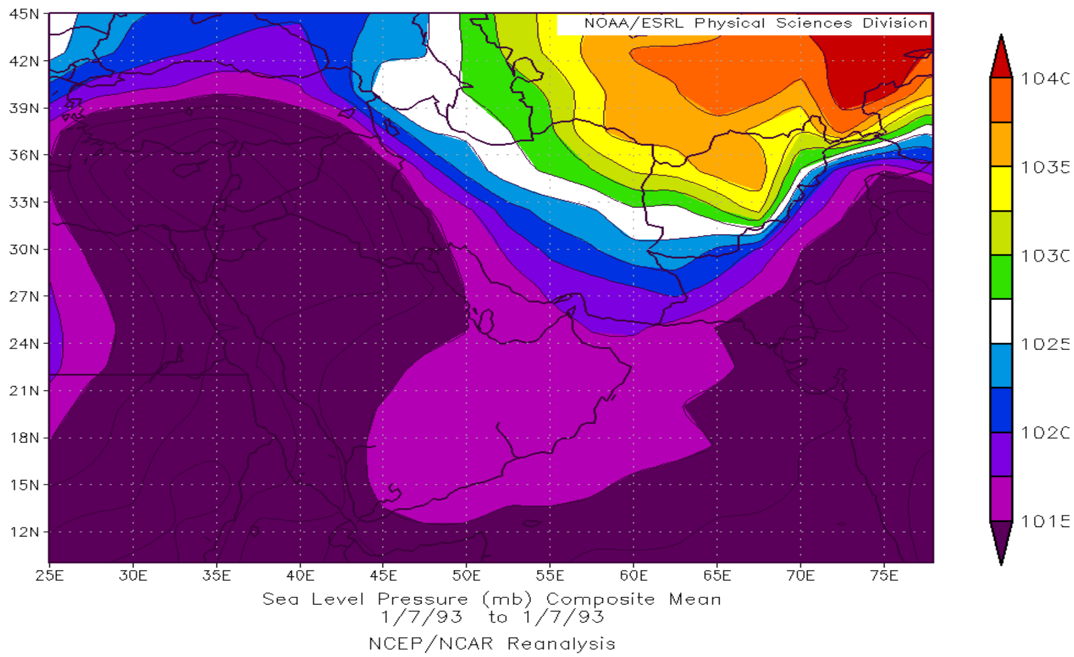

Figure 29. Sea Level Pressure, 7 January 1993.

B. 850-HPA GEOPOTENTIAL HEIGHTS, 1–7 JANUARY 1993

This section shows NCEP/NCAR re-analyses of geopotential heights for 850-hPa pressure levels for the Middle East region in general, and specifically the operating area in the campaign scenario in WIAT.

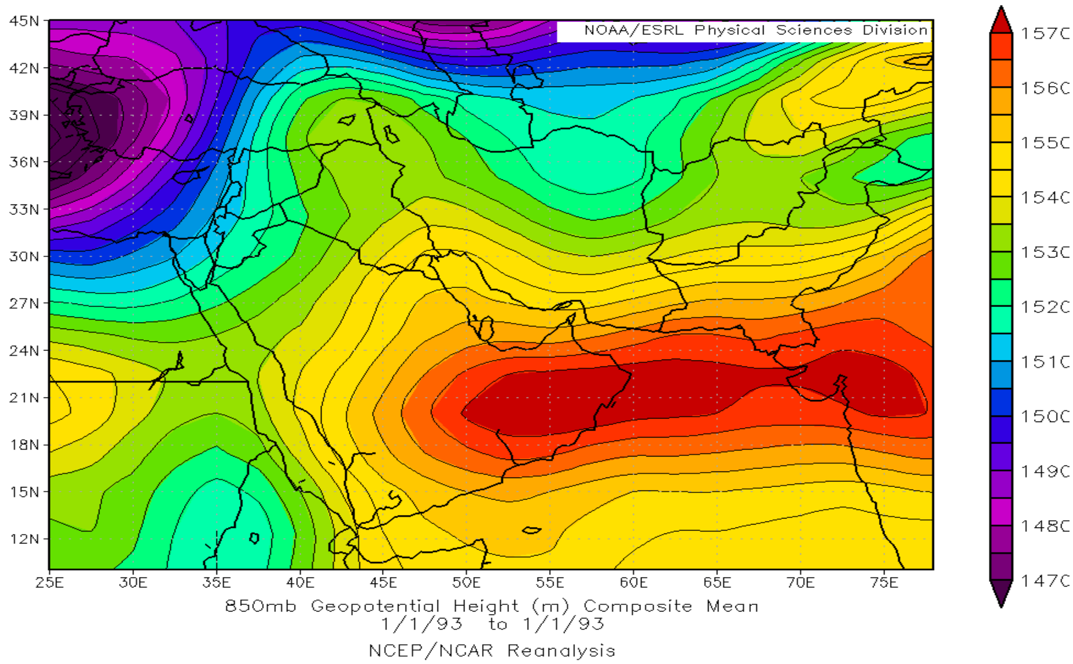

Figure 30. 850 hPa Geopotential Heights, 1 January 1993.

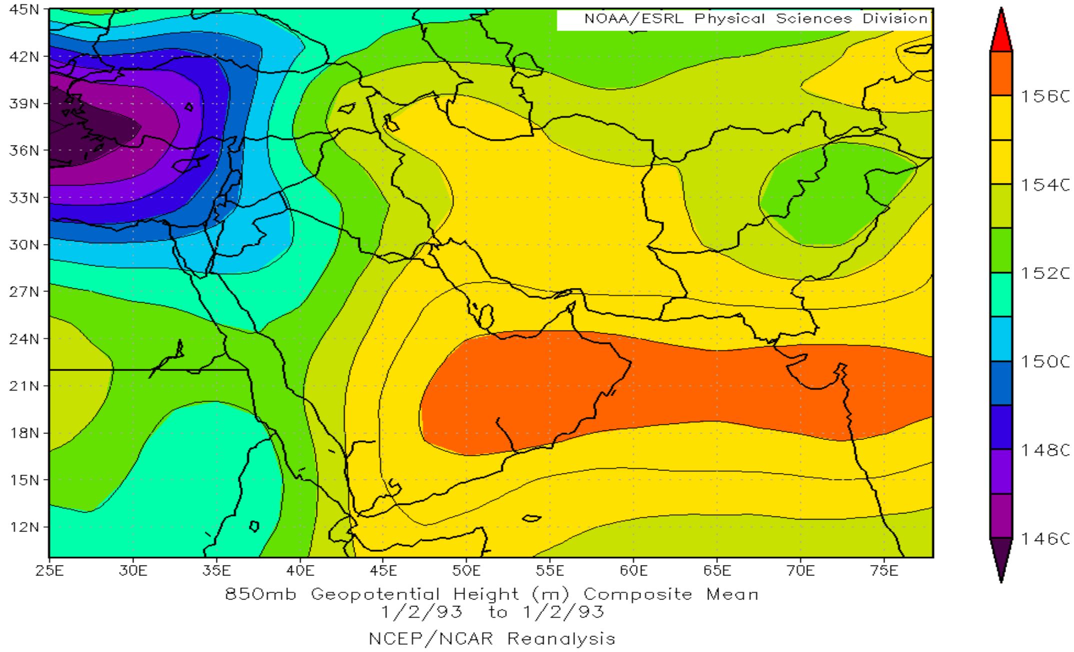

Figure 31. 850 hPa Geopotential Heights, 2 January 1993.

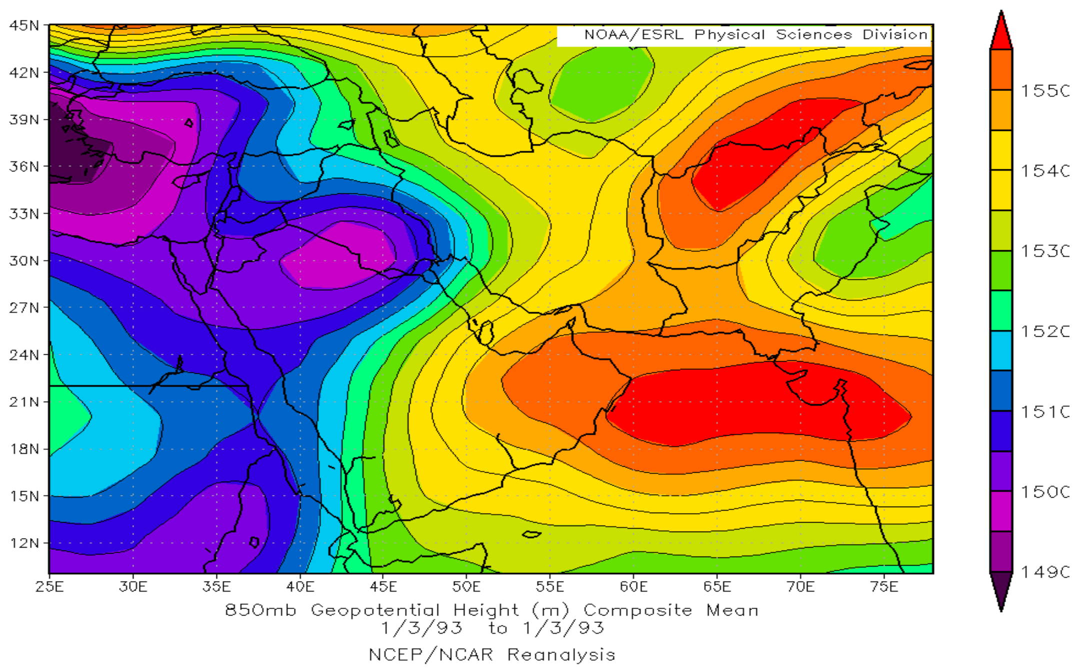

Figure 32. 850 hPa Geopotential Heights, 3 January 1993.

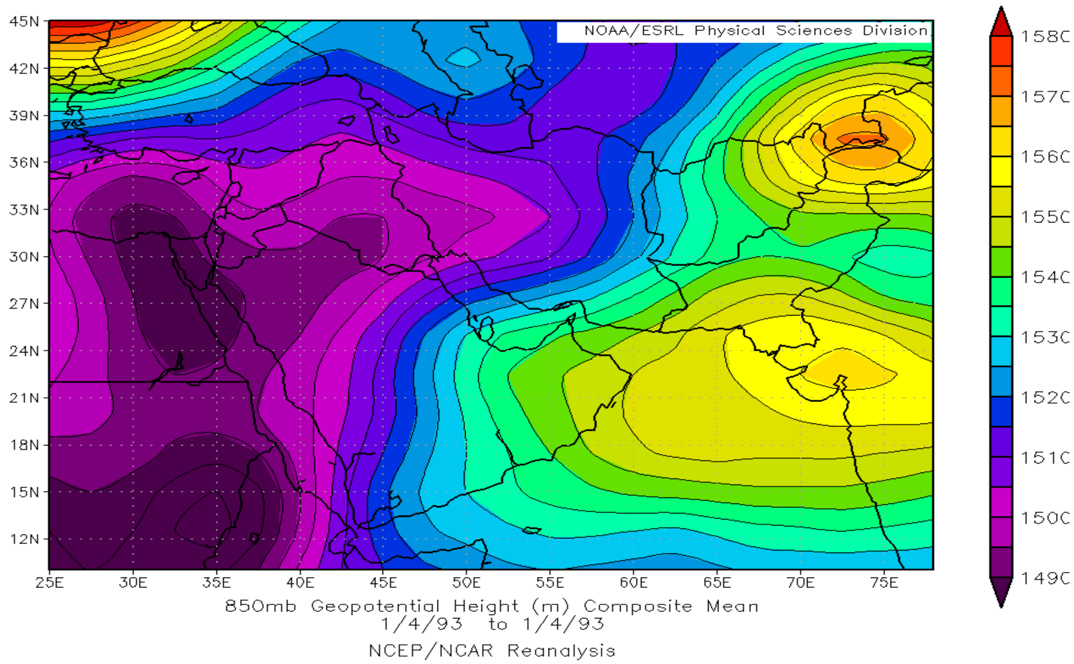

Figure 33. 850 hPa Geopotential Heights, 4 January 1993.

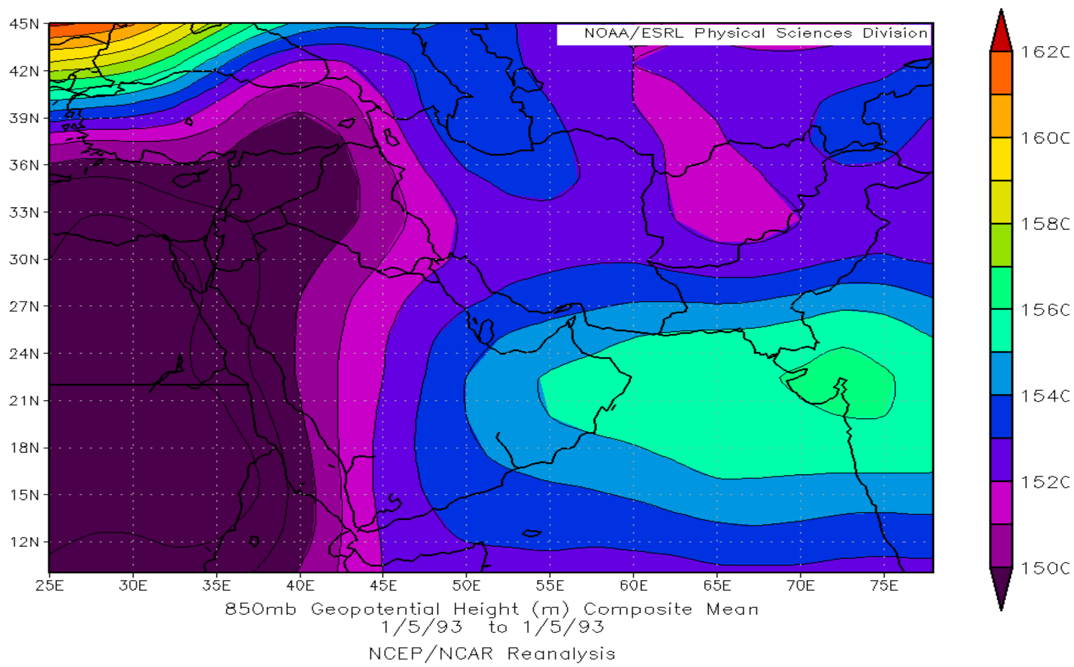

Figure 34. 850 hPa Geopotential Heights, 5 January 1993.

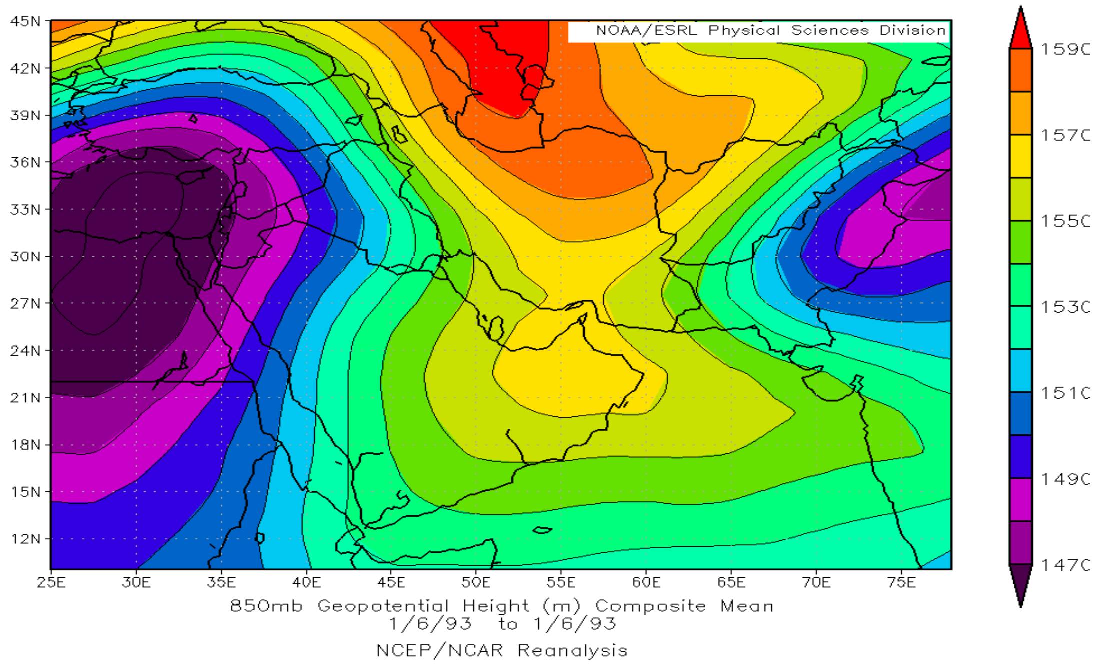

Figure 35. 850 hPa Geopotential Heights, 6 January 1993.

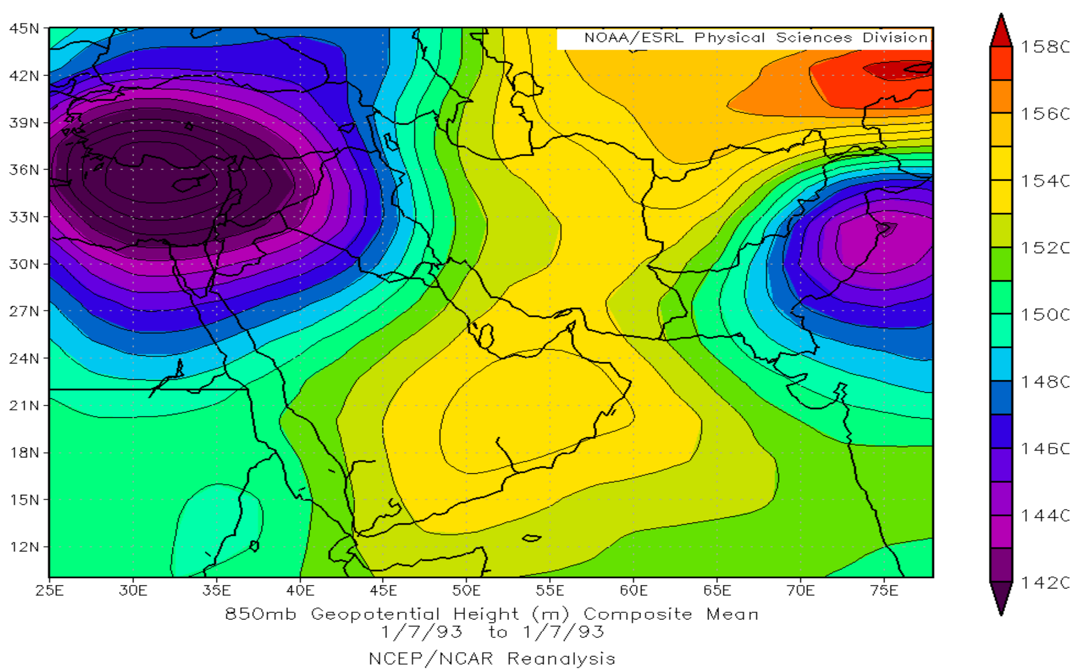

Figure 36. 850 hPa Geopotential Heights, 7 January 1993.

C. 500-HPA GEOPOTENTIAL HEIGHTS, 1–7 JANUARY 1993

This section shows NCEP/NCAR re-analyses of geopotential heights for 500-hPa pressure levels for the Middle East region in general, and specifically the operating area in the campaign scenario in WIAT.

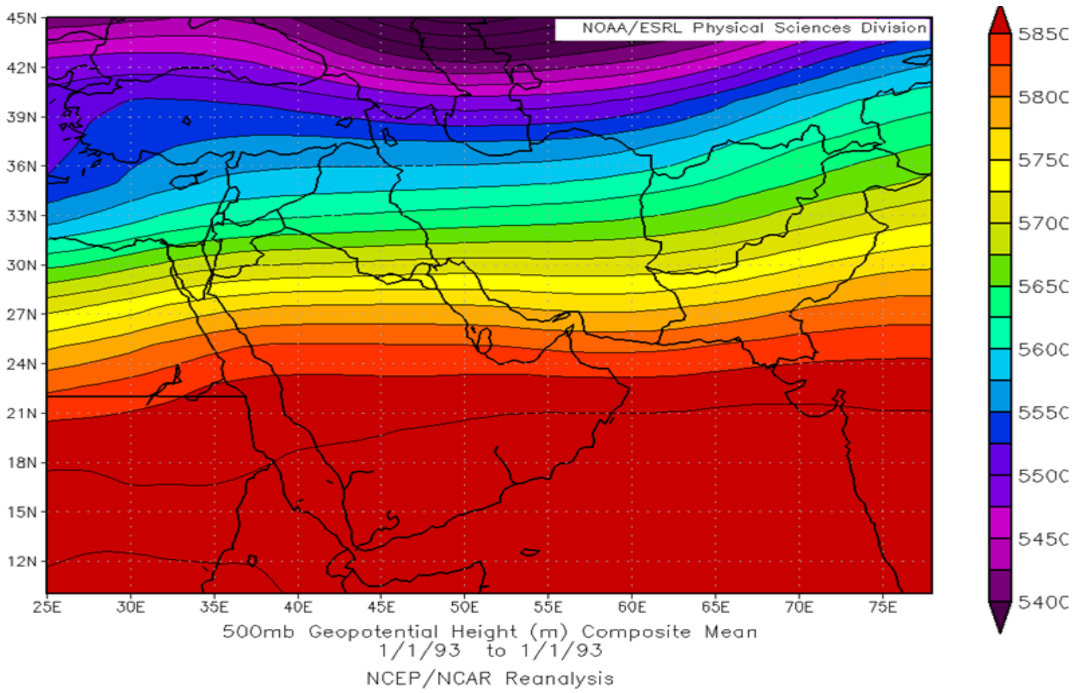

Figure 37. 500 hPa Geopotential Heights, 1 January 1993.

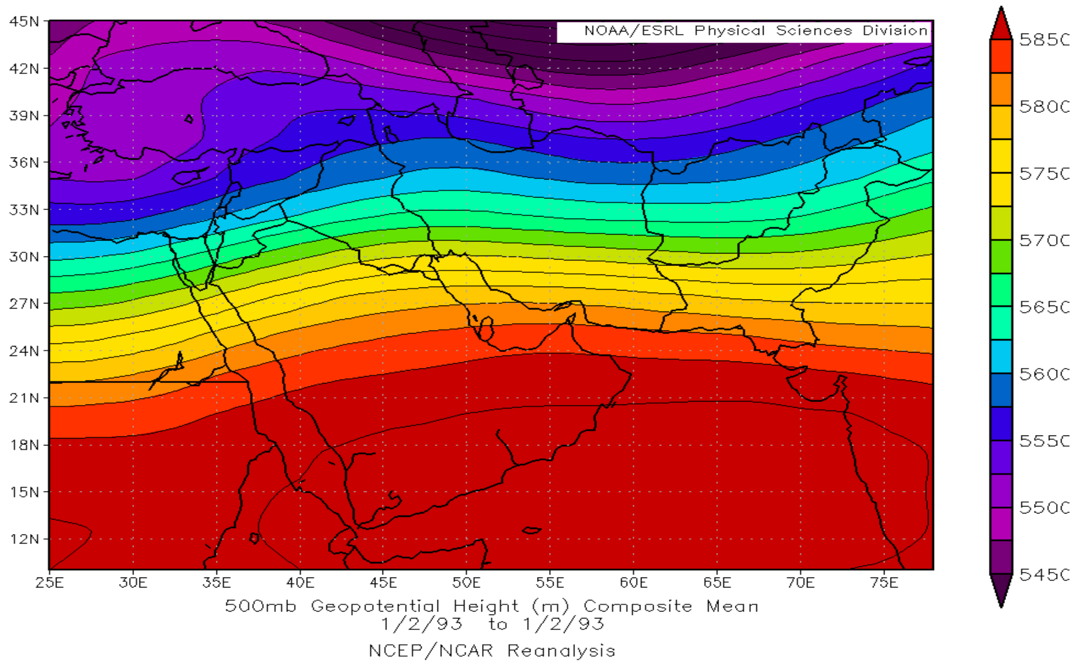

Figure 38. 500 hPa Geopotential Heights, 2 January 1993.

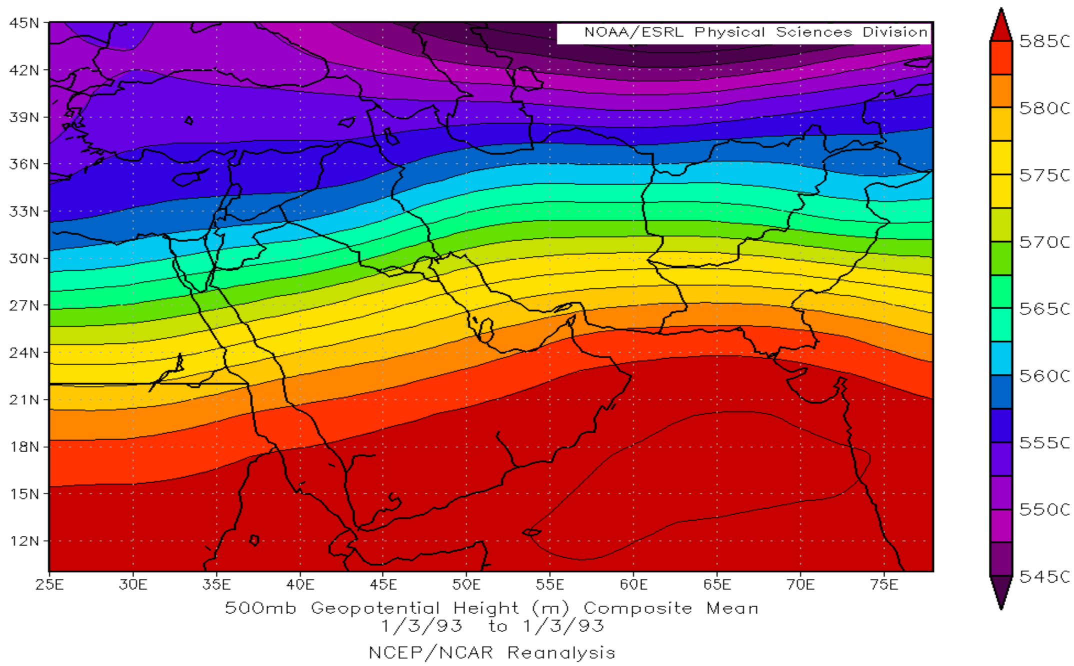

Figure 39. 500 hPa Geopotential Heights, 3 January 1993.

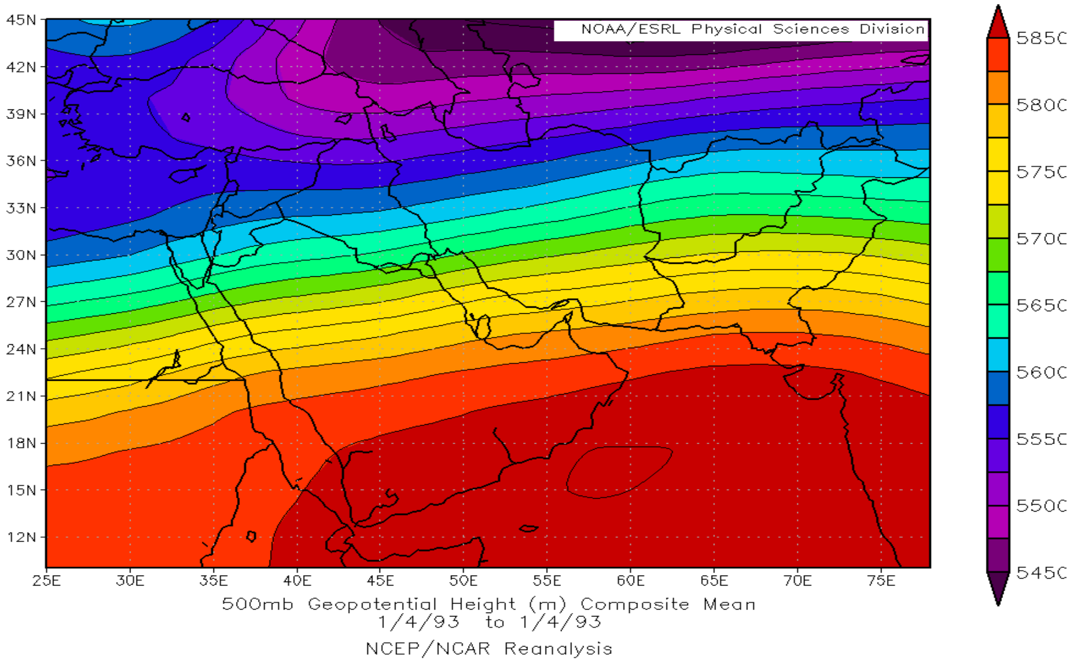

Figure 40. 500 hPa Geopotential Heights, 4 January 1993.

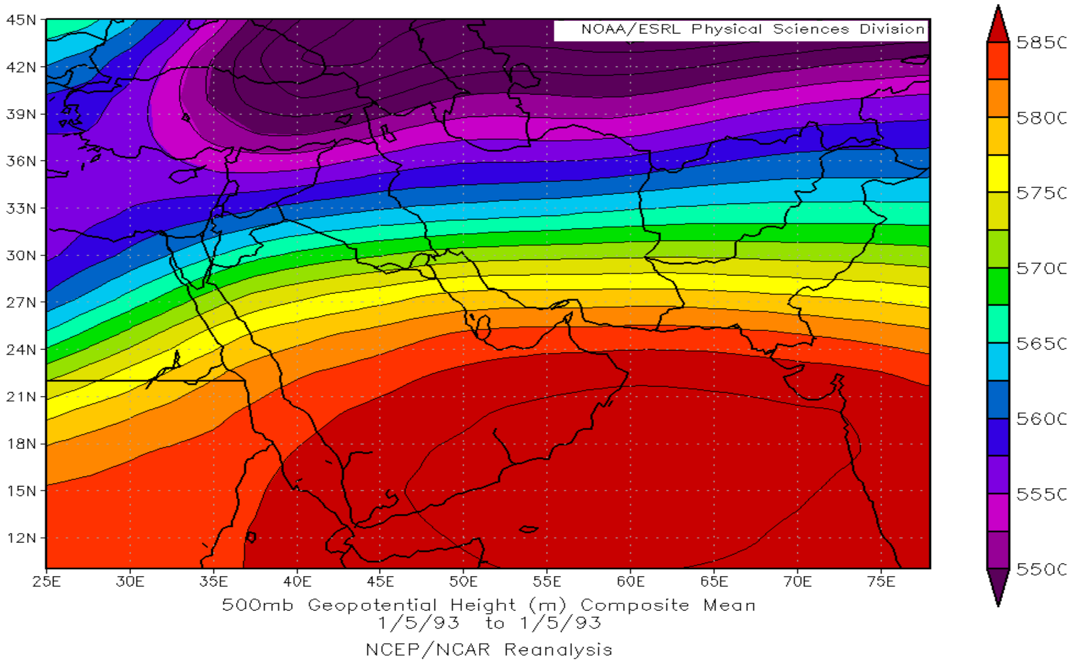

Figure 41. 500 hPa Geopotential Heights, 5 January 1993.

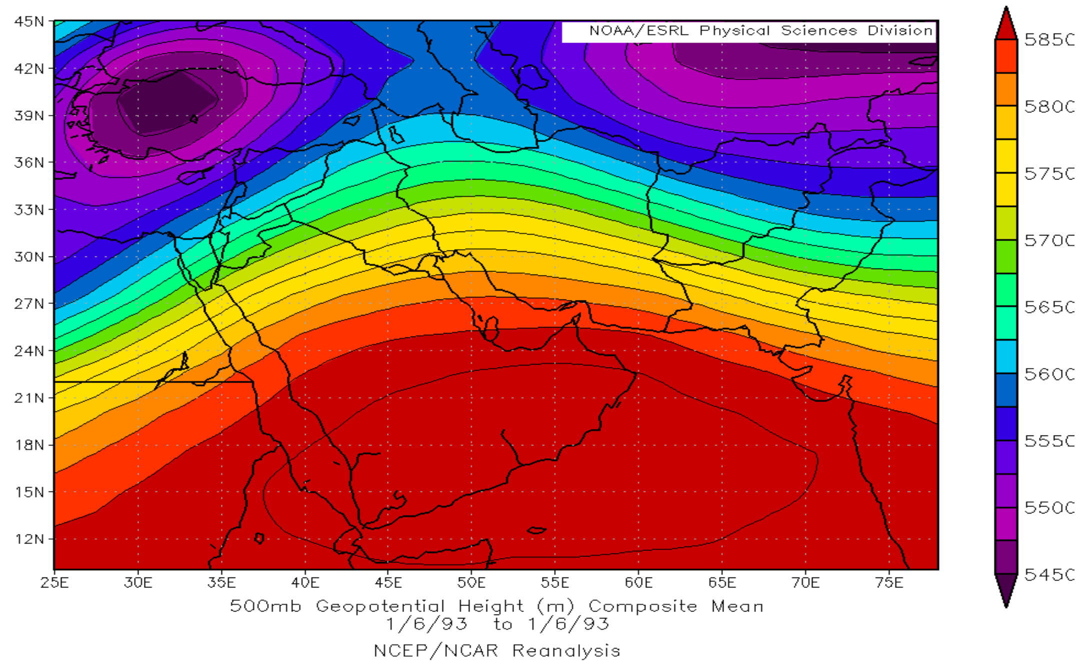

Figure 42. 500 hPa Geopotential Heights, 6 January 1993.

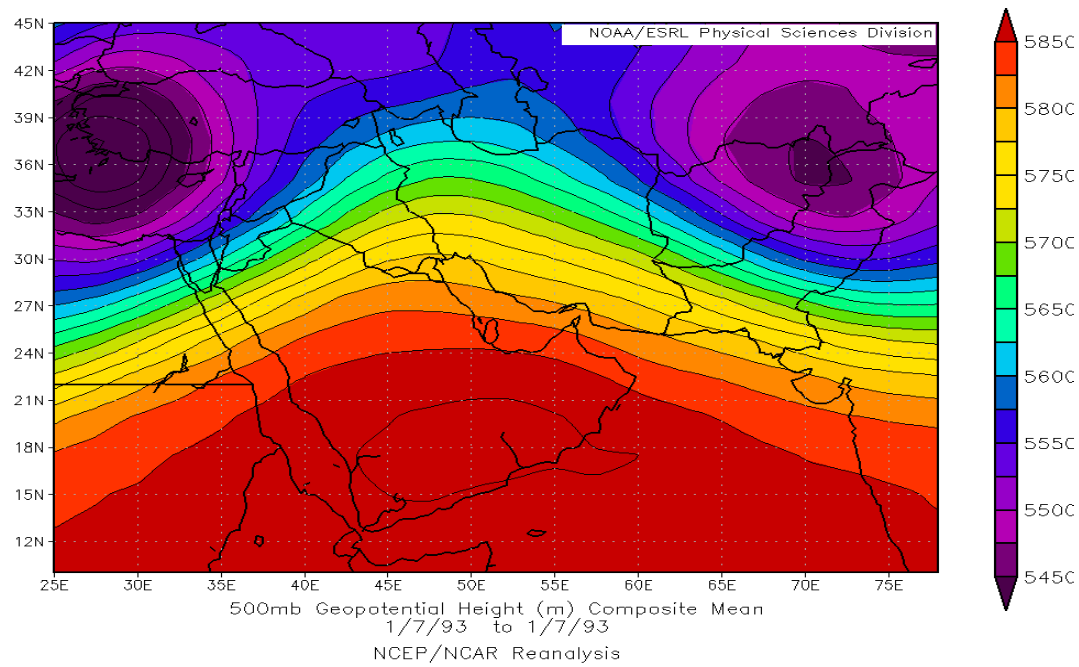

Figure 43. 500 hPa Geopotential Heights, 7 January 1993.

D. SURFACE PRECIPITATION RATE, 1–7 JANUARY 1993

This section shows NCEP/NCAR re-analyses of surface precipitation rate for the Middle East region in general, and specifically the operating area in the campaign scenario in WIAT.

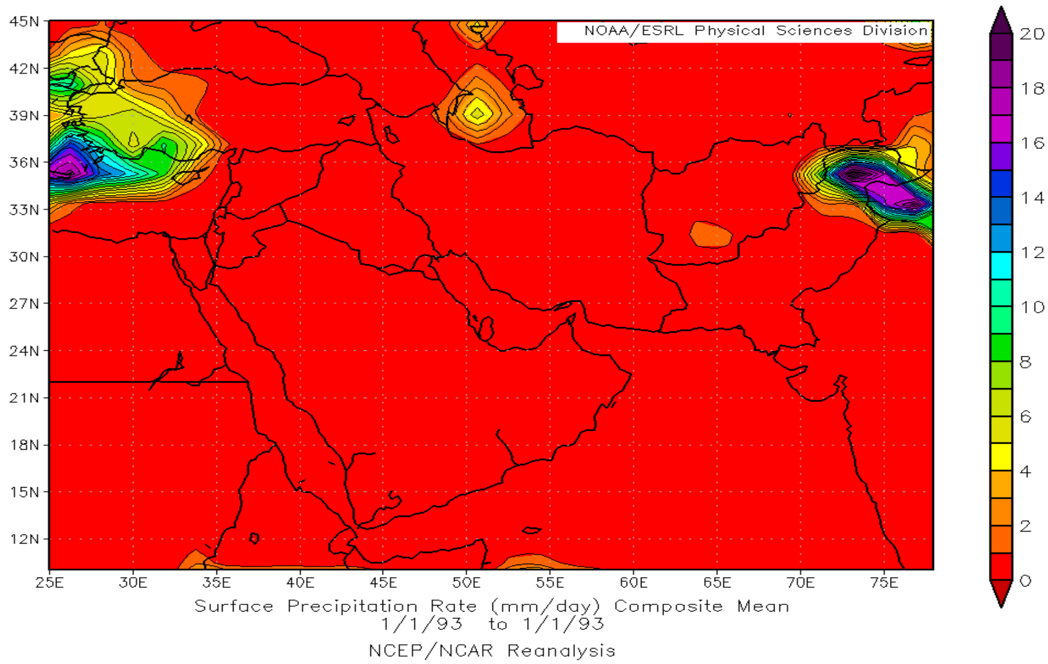

Figure 44. Surface Precipitation Rate, 1 January 1993.

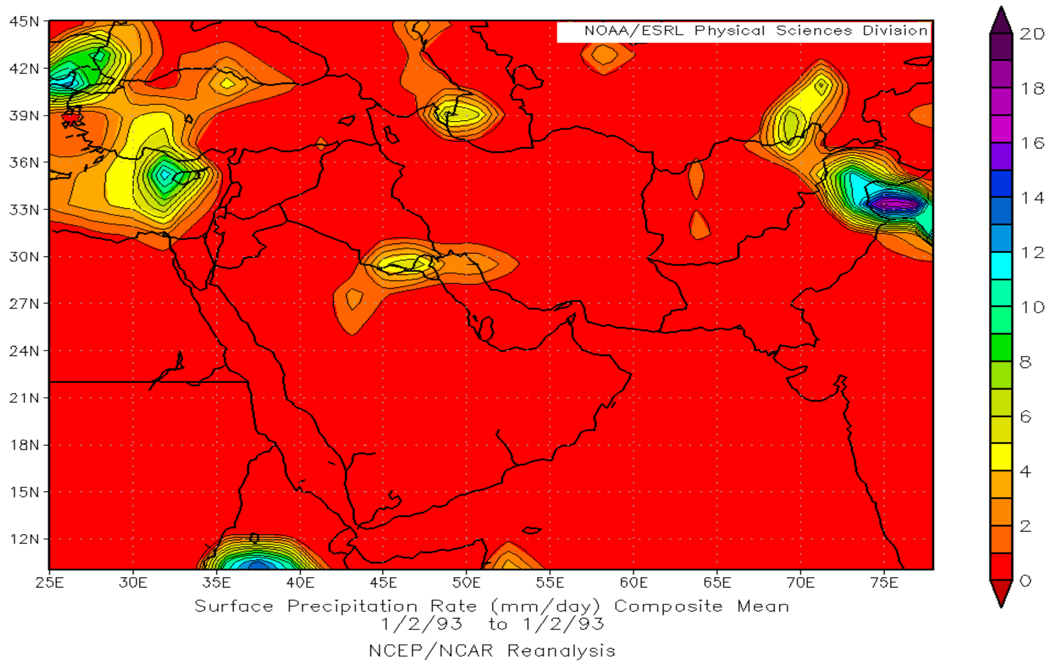

Figure 45. Surface Precipitation Rate, 2 January 1993.

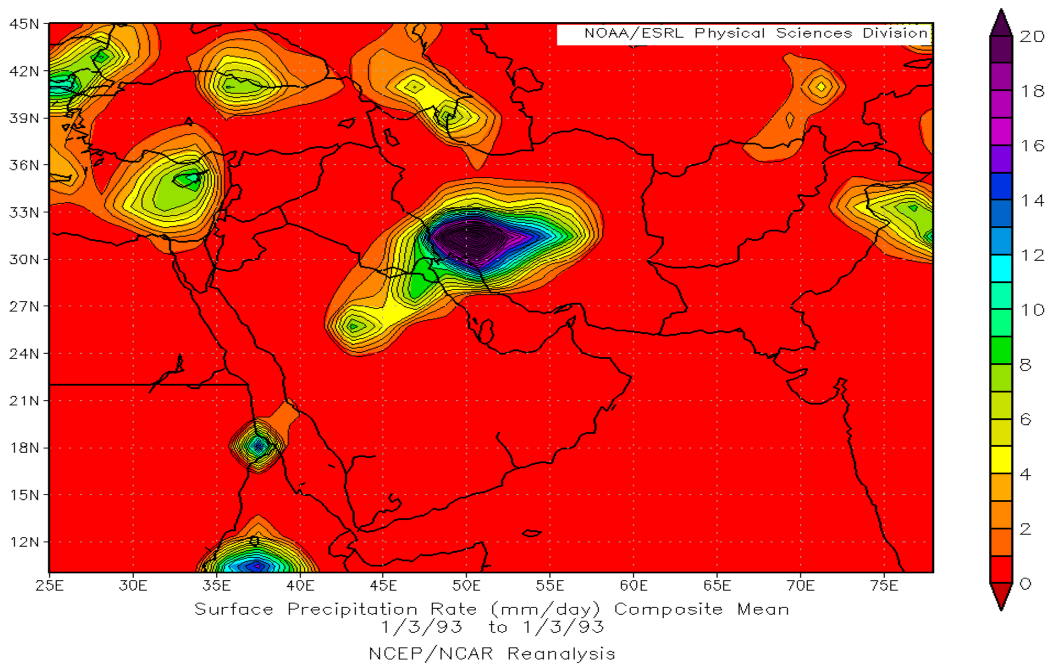

Figure 46. Surface Precipitation Rate, 3 January 1993.

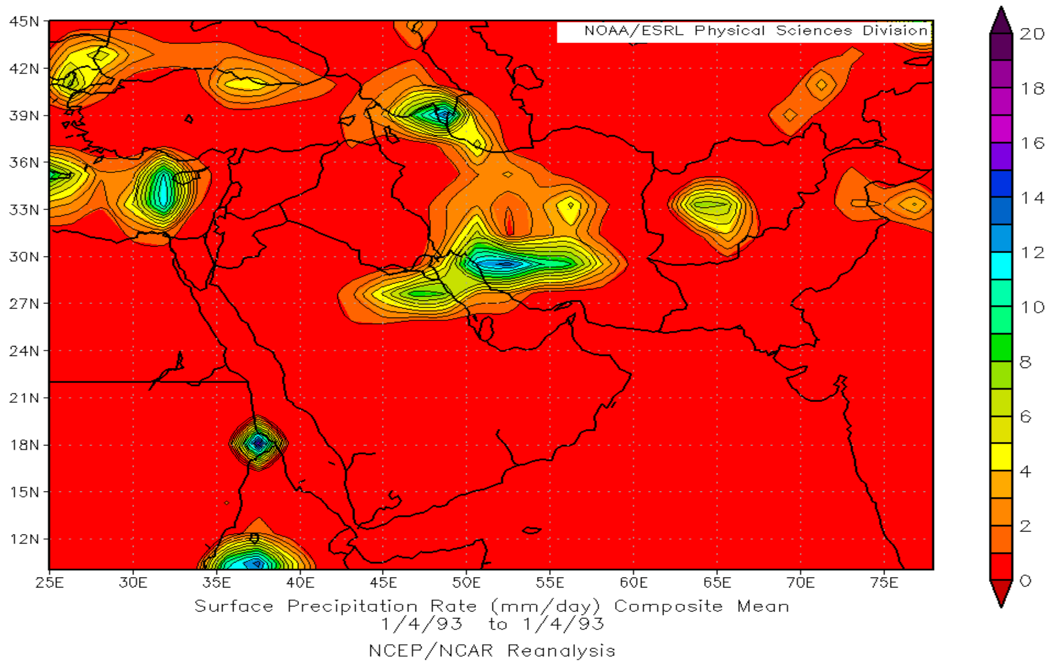

Figure 47. Surface Precipitation Rate, 4 January 1993.

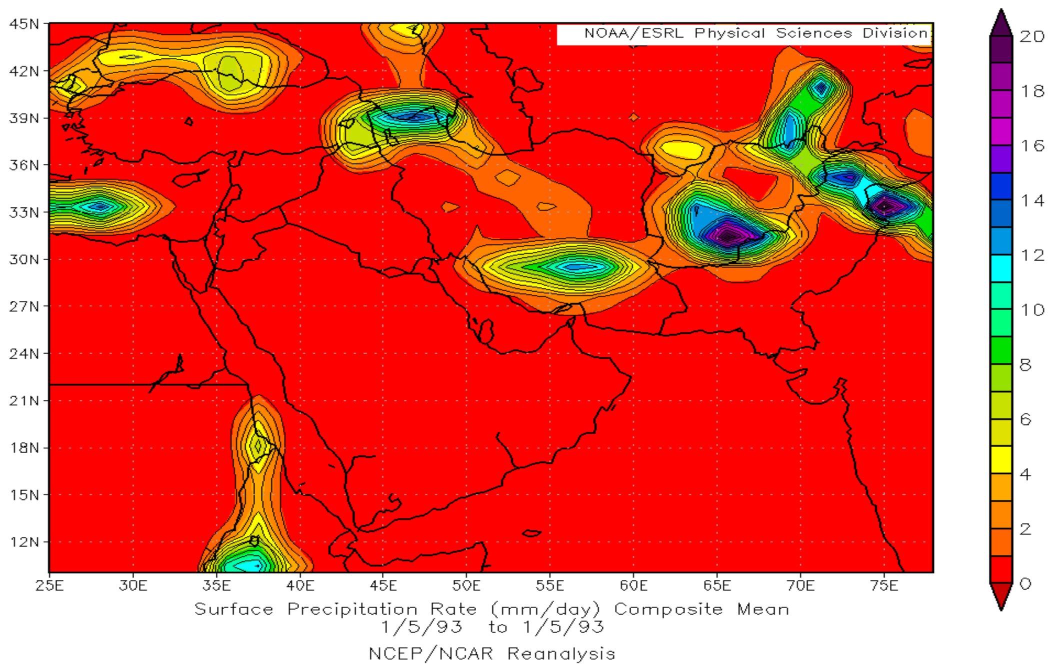

Figure 48. Surface Precipitation Rate, 5 January 1993.

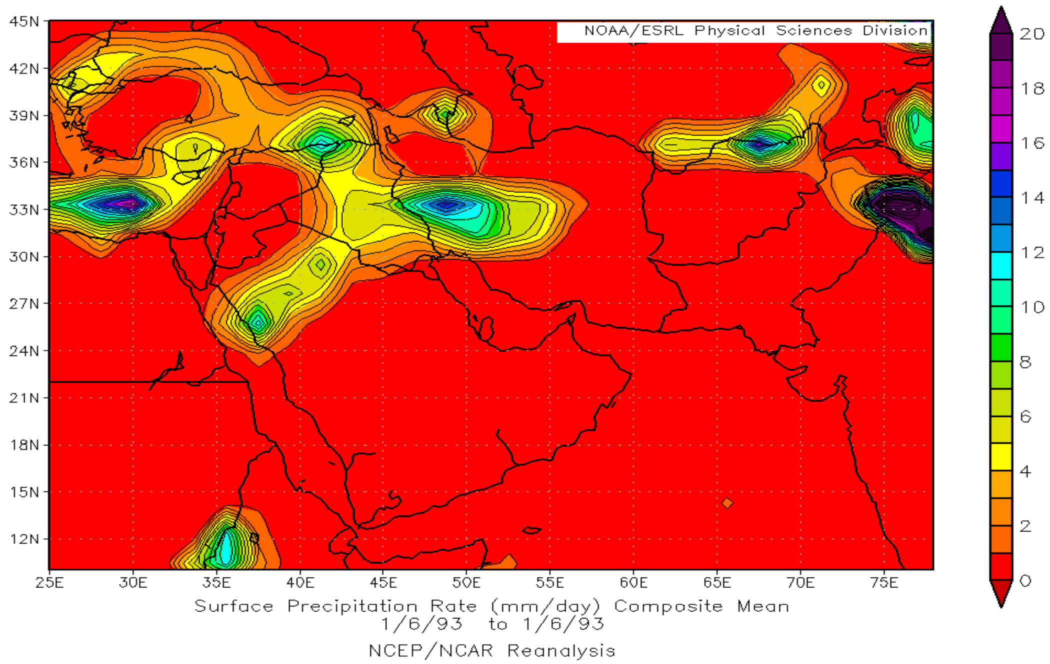

Figure 49. Surface Precipitation Rate, 6 January 1993.

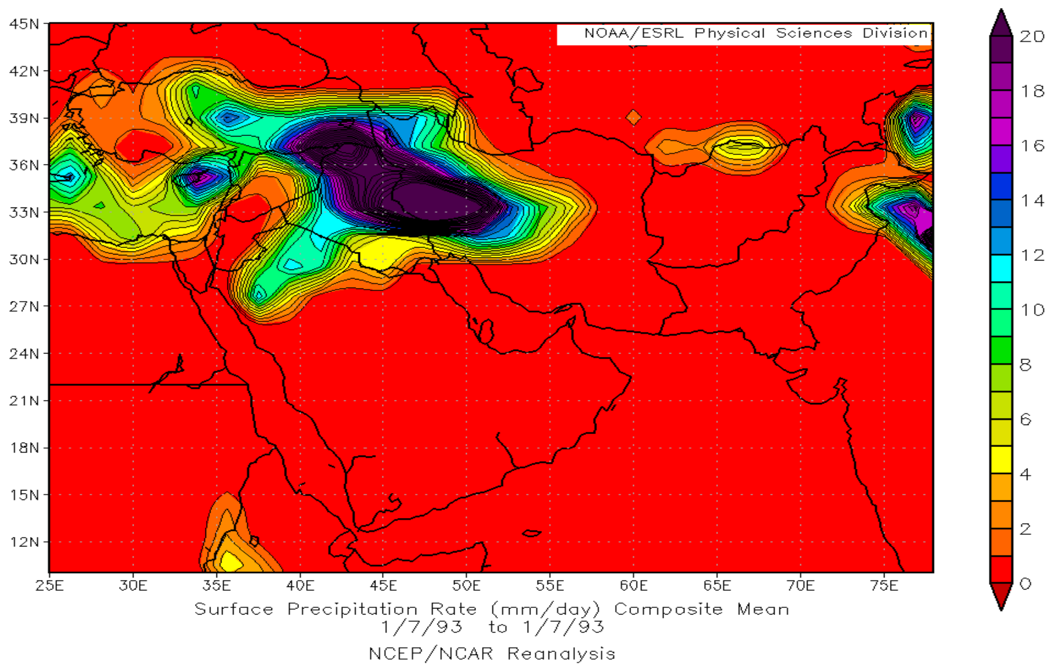

Figure 50. Surface Precipitation Rate, 7 January 1993.

E. SURFACE RELATIVE HUMIDITY, 1–7 JANUARY 1993

This section shows NCEP/NCAR re-analyses of surface relative humidity, the primary factor in determining visibility, for the Middle East region in general, and specifically the operating area in the campaign scenario in WIAT.

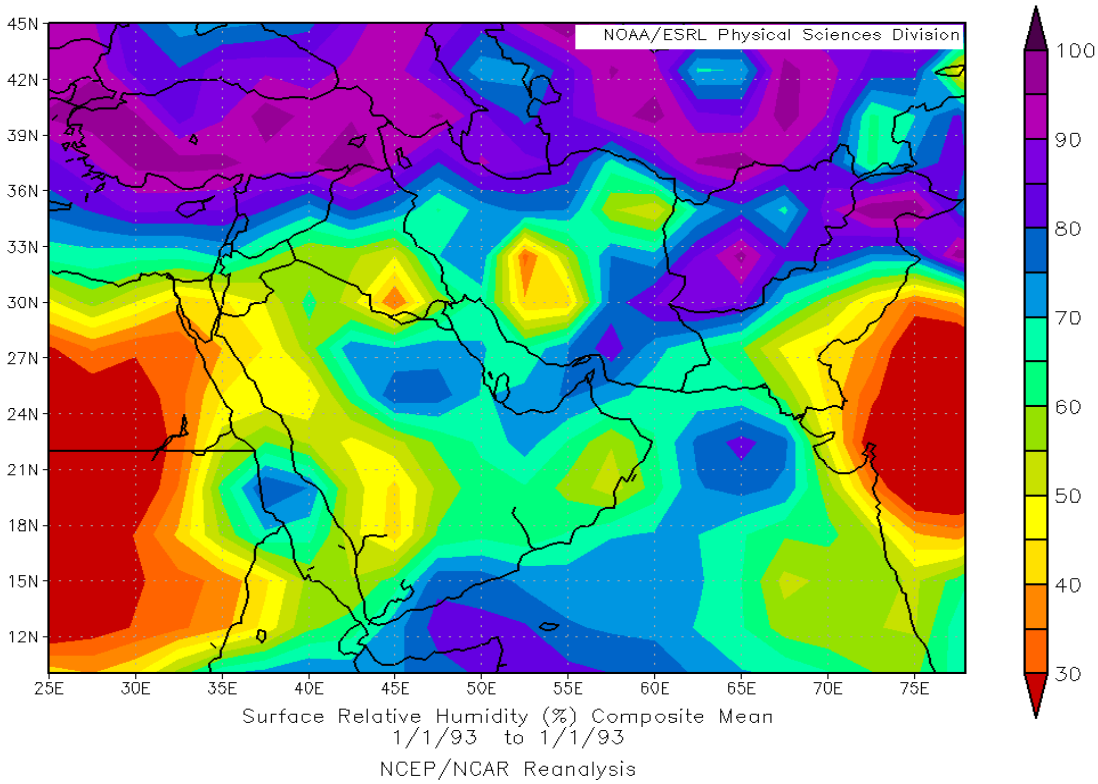

Figure 51. Surface Relative Humidity, 1 January 1993.

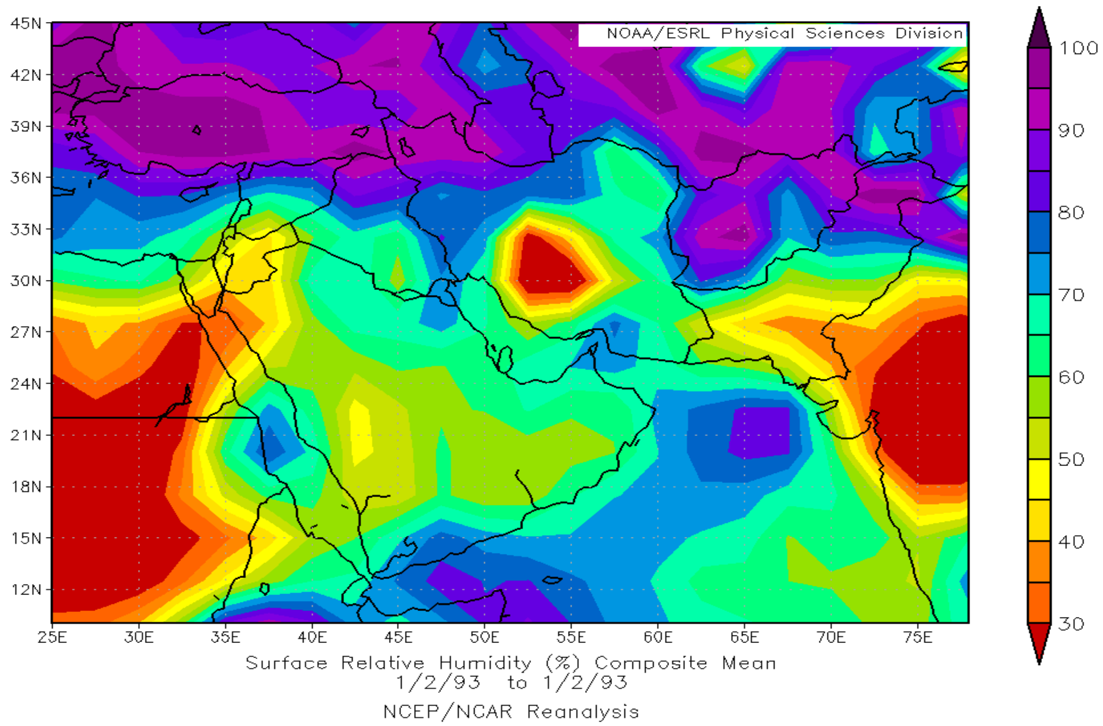

Figure 52. Surface Relative Humidity, 2 January 1993.

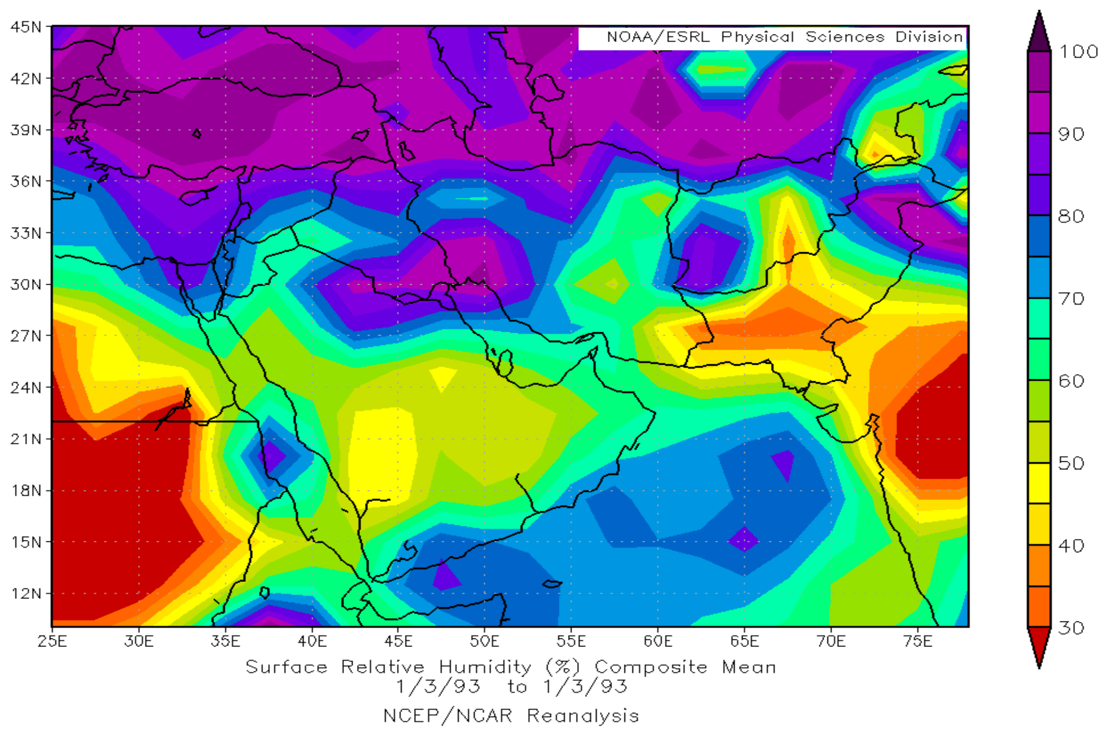

Figure 53. Surface Relative Humidity, 3 January 1993.

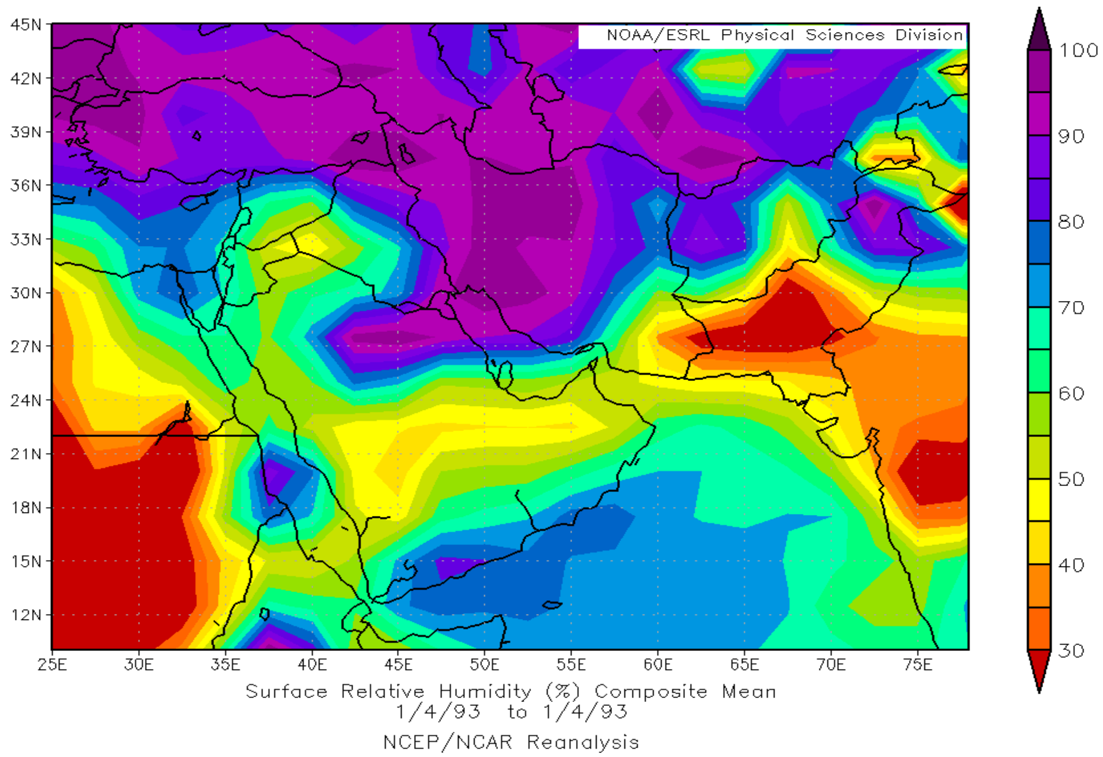

Figure 54. Surface Relative Humidity, 4 January 1993.

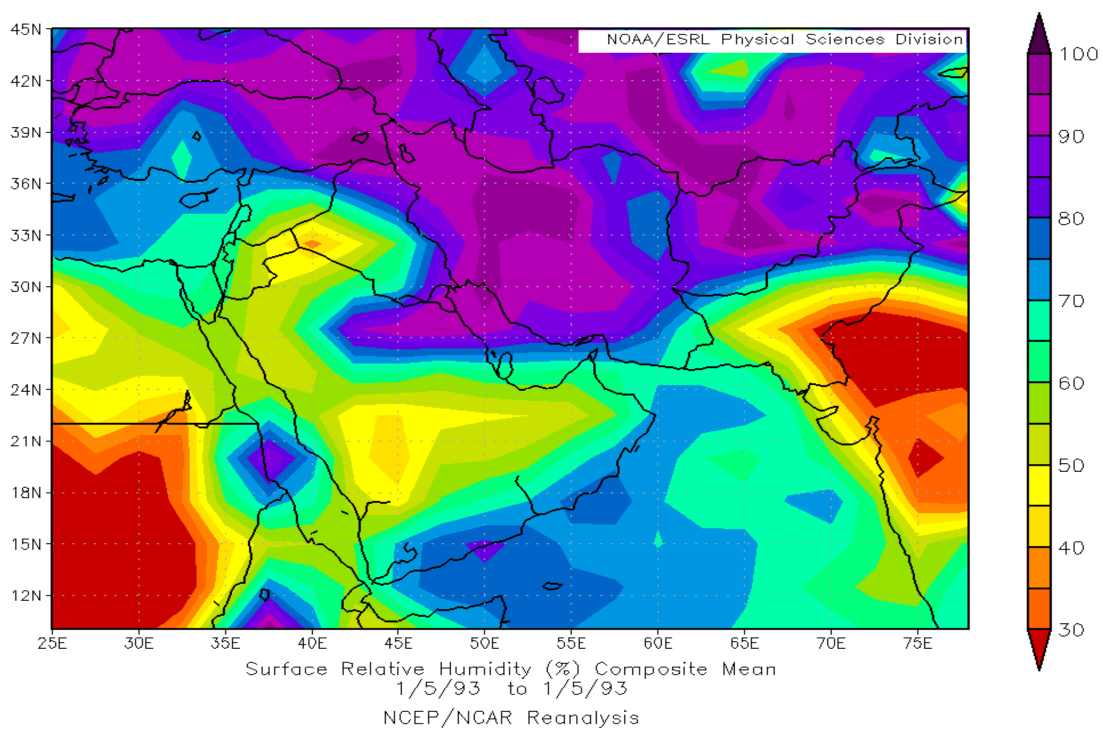

Figure 55. Surface Relative Humidity, 5 January 1993.

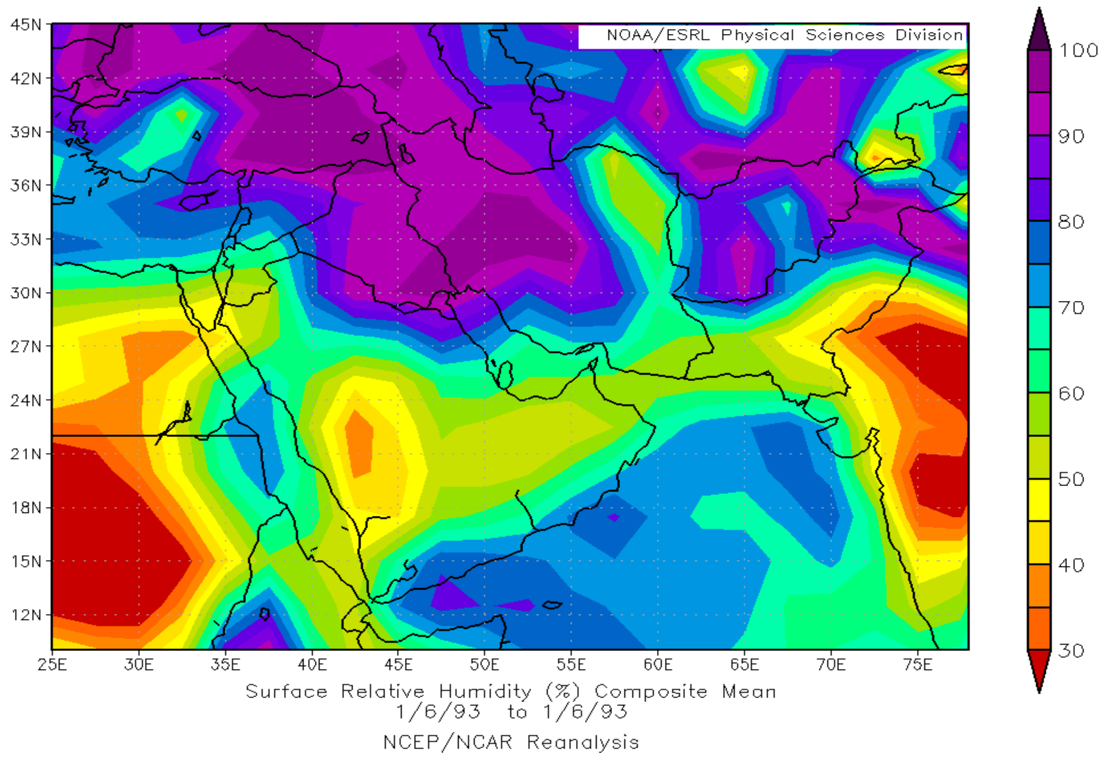

Figure 56. Surface Relative Humidity, 6 January 1993.

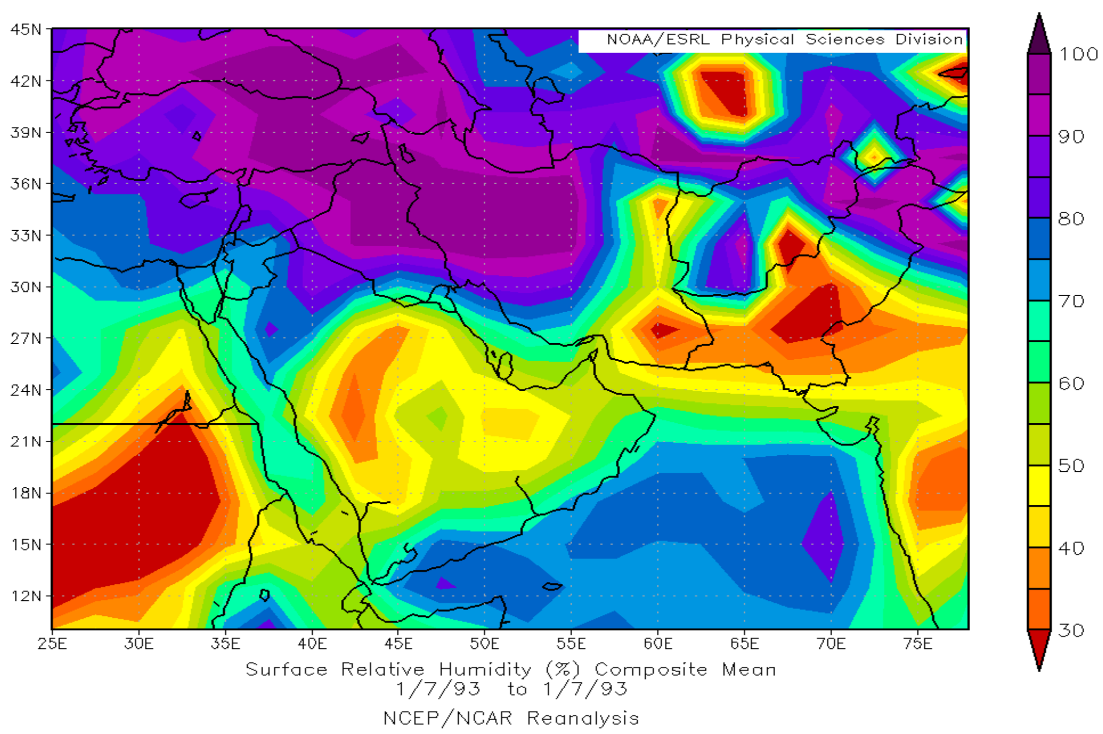

Figure 57. Surface Relative Humidity, 7 January 1993.

APPENDIX B: LIST OF DEFAULT WIAT SETTINGS IN DETERMINISTIC SCENARIOS

This appendix lists the default settings set to use WIAT in a deterministic forecasting scenario.

Weather Prediction Method	1
0 – Perfect Knowledge	
1 – Gaussian Weather Prediction	
2 – Time Shifted Weather Prediction	
3 – Location Shifted Weather Prediction	
4 – Percent Accuracy Weather Prediction	
5 – Hand Entered Forecast Prediction	
Weather Prediction Parameters	
Short Term Parameters	
Sigma for Short Term Forecast	0.5 / 0.75 / 1.0
Short Term Forecast Time Shift	Not applicable
Short Term Vertical Shift	Not applicable
Short Term Horizontal Shift	Not applicable
Short Term Green Percent Accuracy	Not applicable
Short Term Yellow Percent Accuracy	Not applicable
Short Term Red Percent Accuracy	Not applicable
Long Term Parameters	
Sigma for Long Term Forecast	1.0 / 1.5 / 2.0 / 2.5
Long Term Forecast Time Shift	Not applicable
Long Term Vertical Shift	Not applicable
Long Term Horizontal Shift	Not applicable
Long Term Green Percent Accuracy	Not applicable
Long Term Yellow Percent Accuracy	Not applicable
Long Term Red Percent Accuracy	Not applicable
Blue Sensor Environmental Degradation	
Blue Ground Sensor Environmental Degradation Vector	
Green Visibility / Green Precipitation	1.0
Yellow Visibility / Green Precipitation	0.9
Yellow Visibility / Yellow Precipitation	0.8
Red Visibility / Green Precipitation	0.7
Red Visibility / Yellow Precipitation	0.6
Red Visibility / Red Precipitation	0.5
Blue UAV Sensor Environmental Degradation Vector	
Green Visibility / Green Precipitation	1.0

Yellow Visibility / Green Precipitation	0.9
Yellow Visibility / Yellow Precipitation	0.8
Red Visibility / Green Precipitation	0.7
Red Visibility / Yellow Precipitation	0.6
Red Visibility / Red Precipitation	0.5

Blue Fighter Sensor Environmental Degradation Vector

Green Visibility / Green Precipitation	1.0
Yellow Visibility / Green Precipitation	0.9
Yellow Visibility / Yellow Precipitation	0.8
Red Visibility / Green Precipitation	0.7
Red Visibility / Yellow Precipitation	0.6
Red Visibility / Red Precipitation	0.5

Blue Weapon Environmental Degradation

Type 1 Weapon P_k Environmental Degradation Vector

Green/Green	1.0
Yellow/Green	0.9
Yellow/Yellow	0.8
Red/Green	0.7
Red/Yellow	0.6
Red/Red	0.5

Type 2 Weapon P_k Environmental Degradation Vector

Green	1.0
Yellow	0.75
Red	0.5

Type 3 Weapon P_k Environmental Degradation Vector

Green/Green	1.0
Yellow/Green	0.9
Yellow/Yellow	0.8
Red/Green	0.7
Red/Yellow	0.6
Red/Red	0.5

Red Environmental Degradation

Red Air Defense Sensor Environmental Degradation Vector

Green/Green	1.0
Yellow/Green	0.9
Yellow/Yellow	0.8
Red/Green	0.7
Red/Yellow	0.6
Red/Red	0.5

Red Ground Sensor Environmental Degradation Vector	
Green/Green	1.0
Yellow/Green	0.9
Yellow/Yellow	0.8
Red/Green	0.7
Red/Yellow	0.6
Red/Red	0.5
Blue Weapon Parameters	
Type 1 Weapon Max Range (nautical miles)	15
Type 2 Weapon Max Range	15
Type 3 Weapon Max Range	15
Type 1 Weapon Max P_k per missile	0.75
Type 2 Weapon Max P_k per missile	0.75
Type 3 Weapon Max P_k per missile	0.75
Type 1 Weapon Single AC Load	2
Type 2 Weapon Single AC Load	2
Type 3 Weapon Single AC Load	2
Blue Fighter Loiter Speed (knots)	350
Blue Fighter Transit Speed	480
Blue Fighter Max Time on Station (hours)	4
Blue UAV Flight Speed	120
Blue Ground SOA	10
Blue Ground SOA In Contact	5
Blue Fighter Sensor Pd	0.9
Blue UAV Sensor Pd	0.9
Blue Ground Detection Pd	0.9
Threat Parameters	
Threat Parameters for Ground Forces	
Total Major Red Ground Forces	5
Red Ground SOA	7
Red Ground SOA In Contact	0
Red Ground Max Pd	0.7
Mobile Threat Target Speed	8
Threat Parameters for Air Defense	
Red Air Defense Weapon Pk	0.05
Red Air Defense Max Sensor Range	15
Red Air Defense Max Pd	0.7
General Inputs	
Number of Simulations	30
Short Term METOC Acceptance Rate	1
Long Term METOC Acceptance Rate	1
Weather Display Update Interval	6 hours

Weather Threshold Values

Max Green LCD	3
Max Green MCD	3
Max Green HCD	3
Max Green Wind Speed (m/s)	9
Min Green Visibility Value (m)	18,300
Max Green Precipitation Value (mm/hr)	8
Max Yellow LCD	6
Max Yellow MCD	6
Max Yellow HCD	6
Max Yellow Wind Speed (m/s)	11
Min Yellow Visibility Value (m)	9,700
Max Yellow Precipitation Value (mm/hr)	10

Impact Thresholds

Mission Length Impact Threshold	1.25
Maximum Extra Transit Time (hours)	1
Sensor Environmental Degradation Threshold Factor	0.5
BDA Environmental Degradation Threshold Factor	0.5

Strike Mission Weather Thresholds

Short Term (Execution)

Max Strike LCD Short Term	6
Max Strike MCD Short Term	6
Max Strike HCD Short Term	6
Max Strike Wind Speed Short Term (m/s)	11
Min Strike Visibility Short Term (m)	9,700
Max Strike Precipitation Short Term (mm/hr)	10

Long Term (Planning)

Max Strike LCD Long Term	6
Max Strike MCD Long Term	6
Max Strike HCD Long Term	6
Max Strike Wind Speed Long Term (m/s)	11
Min Strike Visibility Long Term (m)	9,700
Max Strike Precipitation Long Term (mm/hr)	10

Kill Box Interdiction (KI) Mission Weather Thresholds

Short Term (Execution)

Max KI LCD Short Term	6
Max KI MCD Short Term	6
Max KI HCD Short Term	6
Max KI Wind Speed Short Term (m/s)	11
Min KI Visibility Short Term (m)	9,700
Max KI Precipitation Short Term (mm/hr)	10

Long Term (Planning)

Max KI LCD Long Term	6
Max KI MCD Long Term	6
Max KI HCD Long Term	6
Max KI Wind Speed Long Term (m/s)	11
Min KI Visibility Long Term (m)	9,700
Max KI Precipitation Long Term (mm/hr)	10

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C: LIST OF CHANGED SETTINGS FOR WIAT* STOCHASTIC SCENARIOS

This appendix lists the settings in WIAT that are adjusted in order to set the scenario for WIAT*. It also lists the scaled parameter values that are employed as the probabilistic mission thresholds in WIAT*.

Weather Prediction Method	5
0 – Perfect Knowledge	
1 – Gaussian Weather Prediction	
2 – Time Shifted Weather Prediction	
3 – Location Shifted Weather Prediction	
4 – Percent Accuracy Weather Prediction	
5 – Hand Entered Forecast Prediction	
Weather Prediction Parameters	
Short Term Parameters	
Sigma for Short Term Forecast	Not applicable
Long Term Parameters	
Sigma for Long Term Forecast	Not applicable
Strike Mission Weather Thresholds	
Short Term (Execution)	
Max Strike LCD Short Term	As Shown in Chart
Max Strike MCD Short Term	“
Max Strike HCD Short Term	“
Max Strike Wind Speed Short Term (m/s)	“
Min Strike Visibility Short Term (m)	“
Max Strike Precipitation Short Term (mm/hr)	“
Long Term (Planning)	
Max Strike LCD Long Term	As Shown in Chart
Max Strike MCD Long Term	“
Max Strike HCD Long Term	“
Max Strike Wind Speed Long Term (m/s)	“
Min Strike Visibility Long Term (m)	“
Max Strike Precipitation Long Term (mm/hr)	“
Kill Box Interdiction (KI) Mission Weather Thresholds	
Short Term (Execution)	
Max KI LCD Short Term	As Shown in Chart
Max KI MCD Short Term	“

Max KI HCD Short Term	"
Max KI Wind Speed Short Term (m/s)	"
Min KI Visibility Short Term (m)	"
Max KI Precipitation Short Term (mm/hr)	"

Long Term (Planning)

Max KI LCD Long Term	As Shown in Chart
Max KI MCD Long Term	"
Max KI HCD Long Term	"
Max KI Wind Speed Long Term (m/s)	"
Min KI Visibility Long Term (m)	"
Max KI Precipitation Long Term (mm/hr)	"

Probability Thresholds and Corresponding WIAT Mission Threshold Inputs

	LCD	MCD	HCD	Wind Speed	Visibility	Precipitation
100.0%	8	8	8	14.0338	0.00	0.5000
87.5%	7	7	7	12.2796	8738.87	0.4375
75.0%	6	6	6	10.5254	17477.75	0.3750
62.5%	5	5	5	8.7711	26216.62	0.3125
50.0%	4	4	4	7.0169	34955.50	0.2500
37.5%	3	3	3	5.2627	43694.37	0.1875
25.0%	2	2	2	3.5085	52433.24	0.1250
12.5%	1	1	1	1.7542	61172.12	0.0625

Table 10. Probability thresholds and their corresponding scaled values for input into WIAT* simulations. Also Table 2.

APPENDIX D: DISTRIBUTION OF WEATHER VALUES

This appendix shows figures representing distribution of actual values, as well as average execution and planning forecast values for the 30 ensemble members for each of the low and high uncertainty scenarios. Additional figures are used to demonstrate the effects of binning by using a rounding function designed to counter WIAT's truncation of forecast cloud cover values.

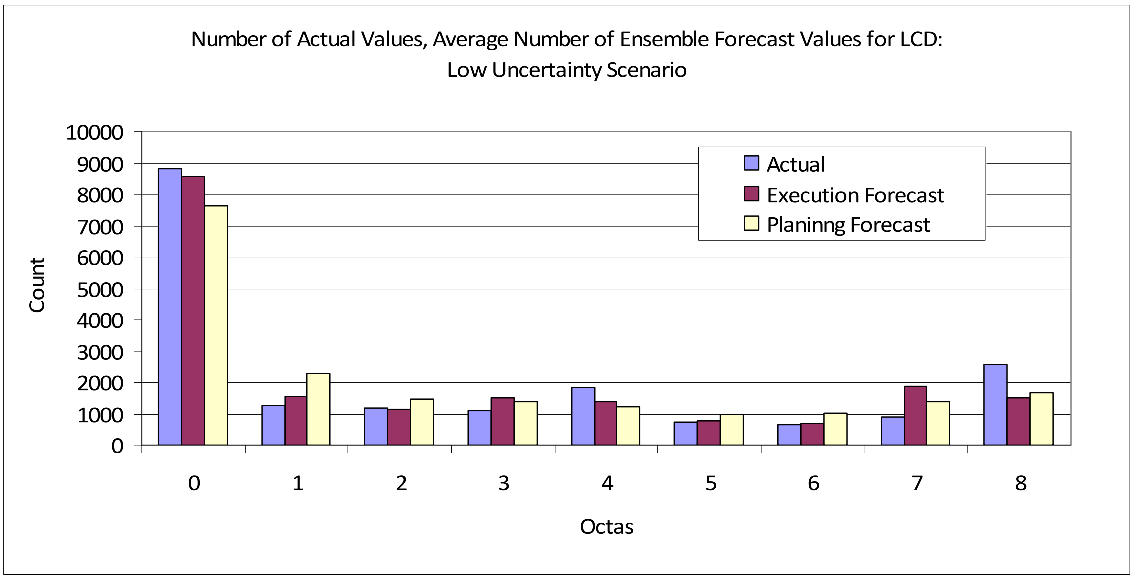

Figure 58. Number of Actual Values and Average Number of Ensemble Forecast Values for LCD: Low Uncertainty Scenario. Values represent the distribution prior to applying the fix for WIAT's truncation of forecast values. Also Figure 15.

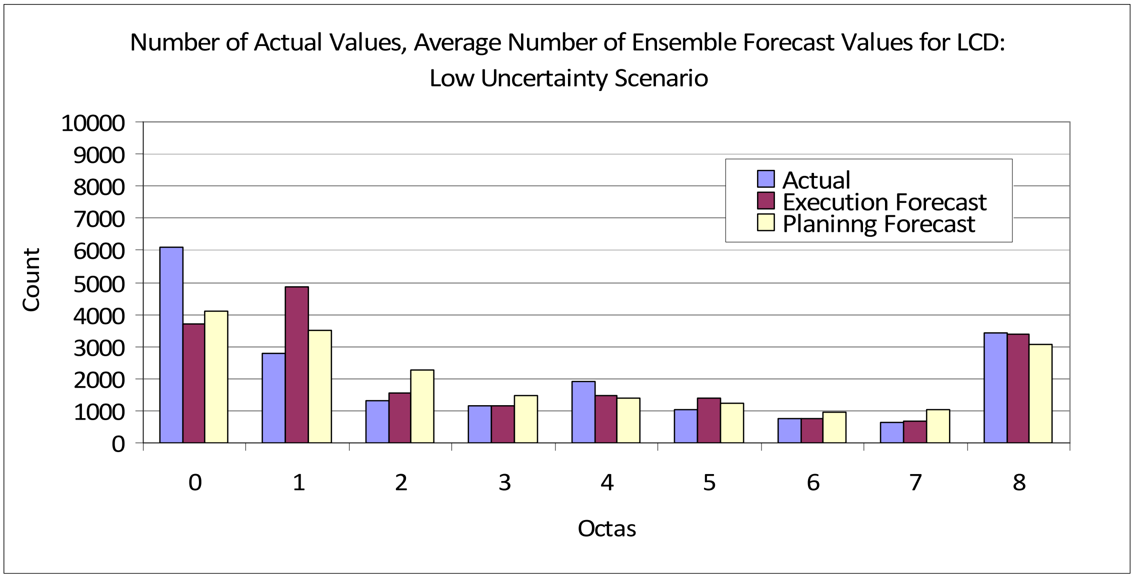

Figure 59. Number of Actual Values and Average Number of Ensemble Forecast Values for LCD: Low Uncertainty Scenario. Values represent the distribution after applying the fix for WIAT's truncation of forecast values.

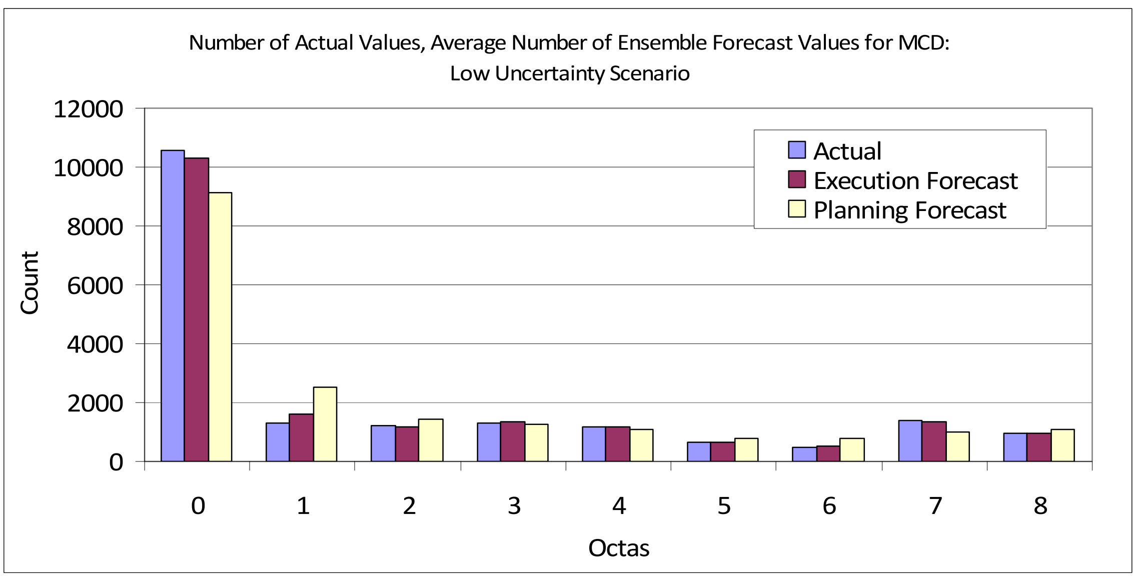

Figure 60. Number of Actual Values and Average Number of Ensemble Forecast Values for MCD: Low Uncertainty Scenario. Values represent the distribution prior to applying the fix for WIAT's truncation of forecast values.

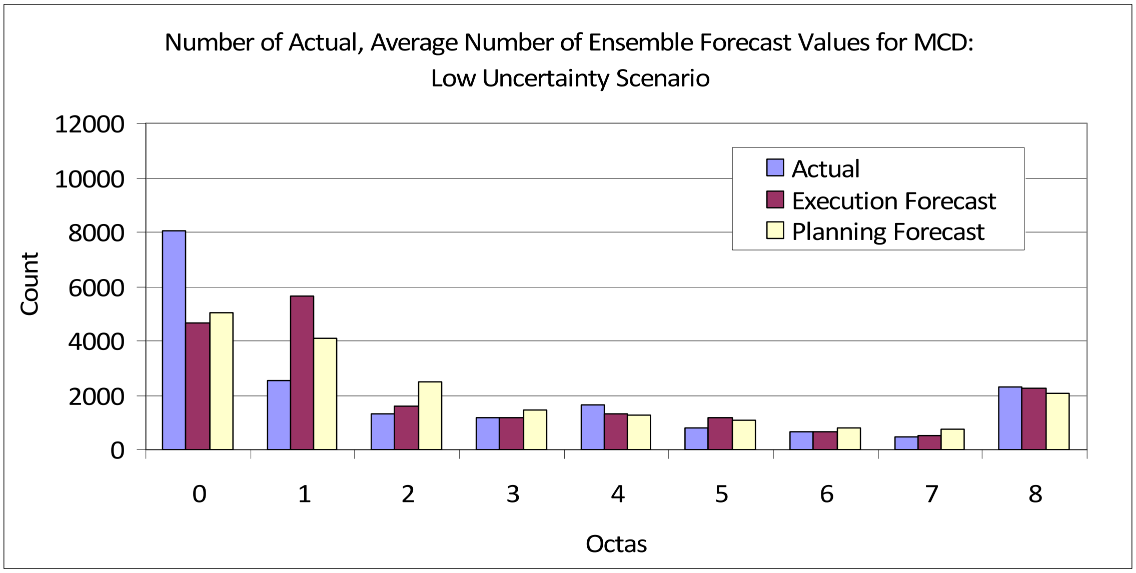

Figure 61. Number of Actual Values and Average Number of Ensemble Forecast Values for MCD: Low Uncertainty Scenario. Values represent the distribution after applying the fix for WIAT's truncation of forecast values.

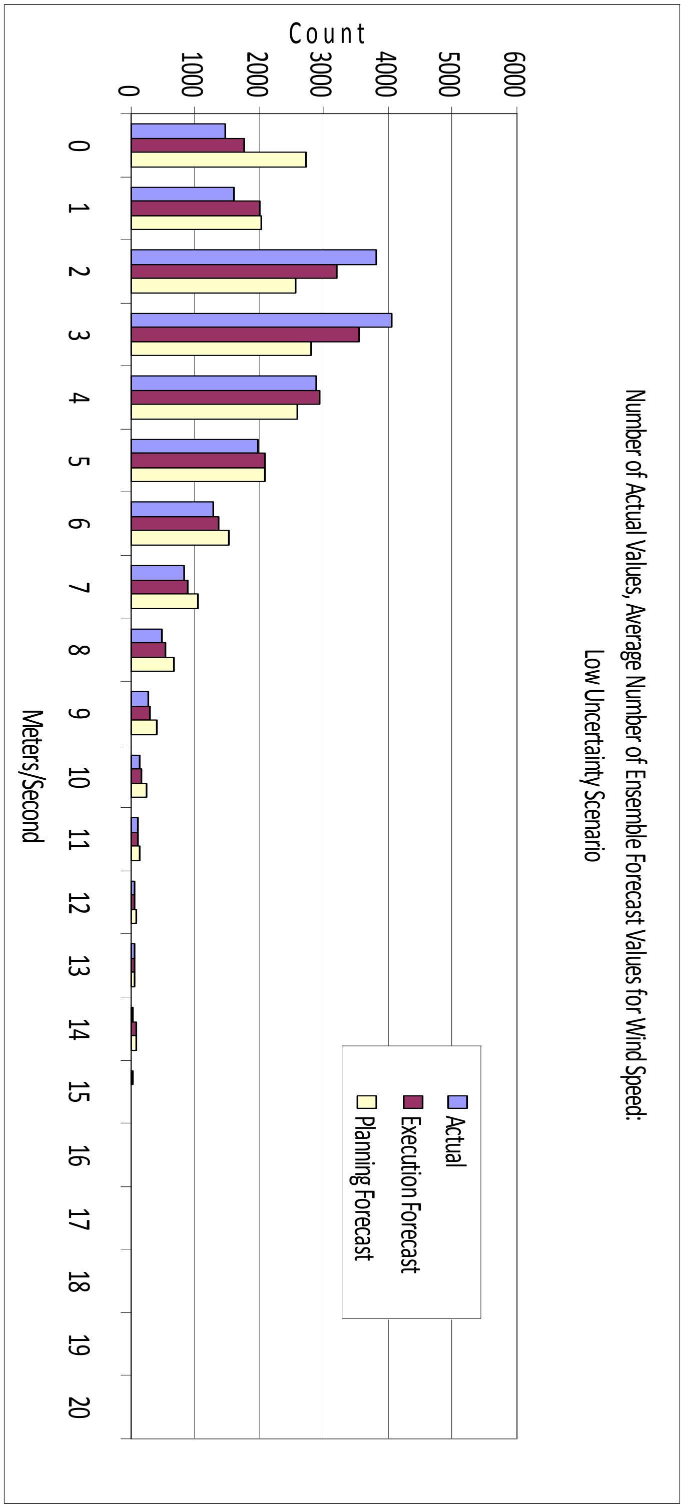

Figure 62. Number of Actual Values and Average Number of Ensemble Forecast Values for Wind Speed: Low Uncertainty Scenario. Also Figure 17.

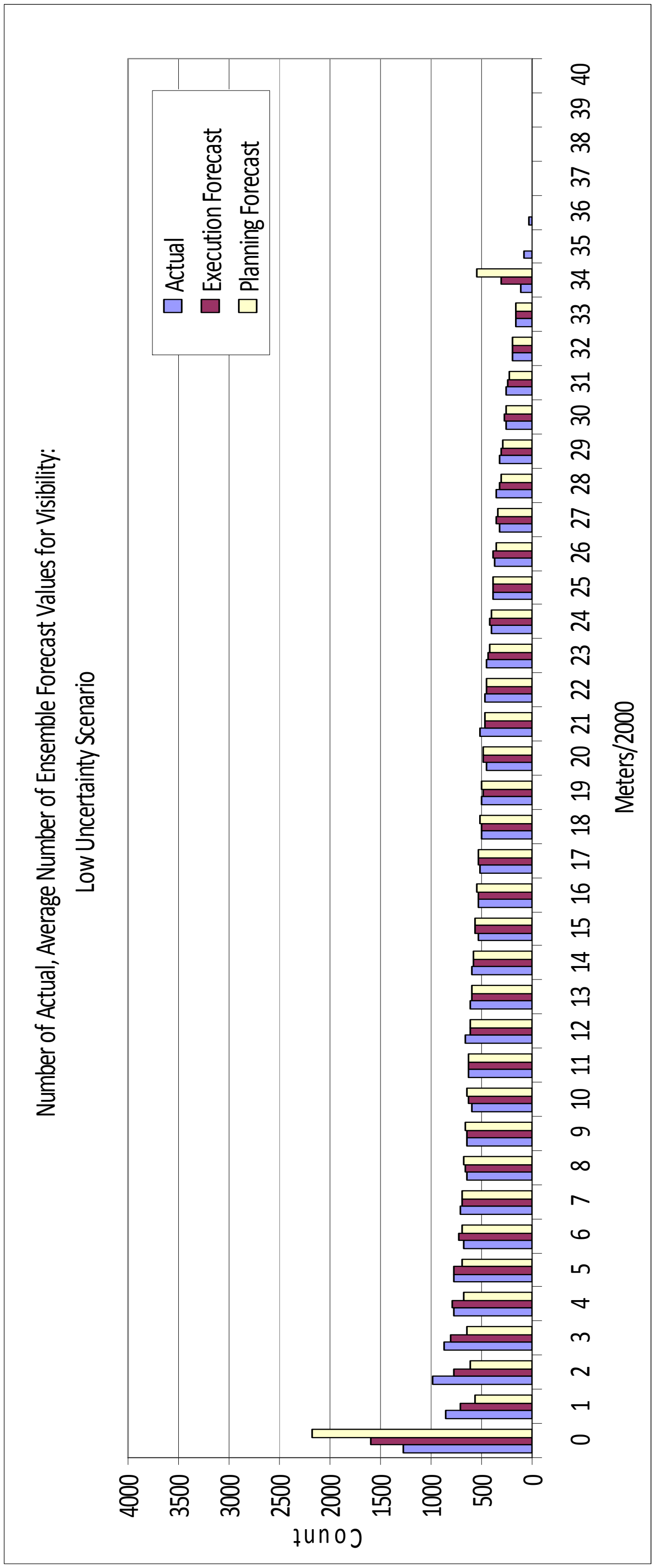

Figure 63. Number of Actual Values and Average Number of Ensemble Forecast Values for Visibility: Low Uncertainty Scenario.

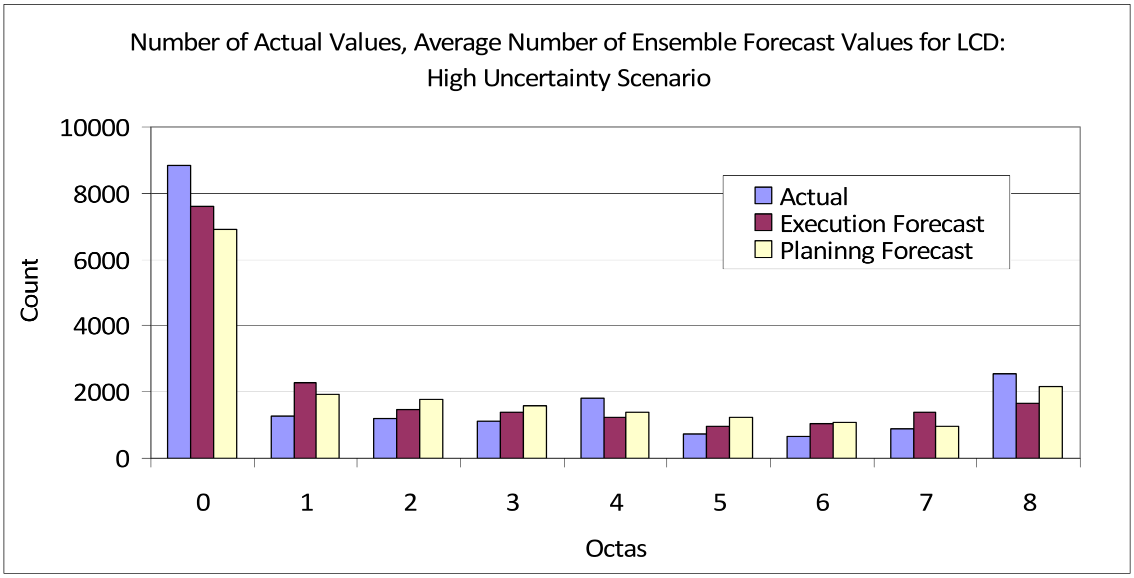

Figure 64. Number of Actual Values and Average Number of Ensemble Forecast Values for LCD: High Uncertainty Scenario. Values represent the distribution prior to applying the fix for WIAT's truncation of forecast values. Also Figure 16.

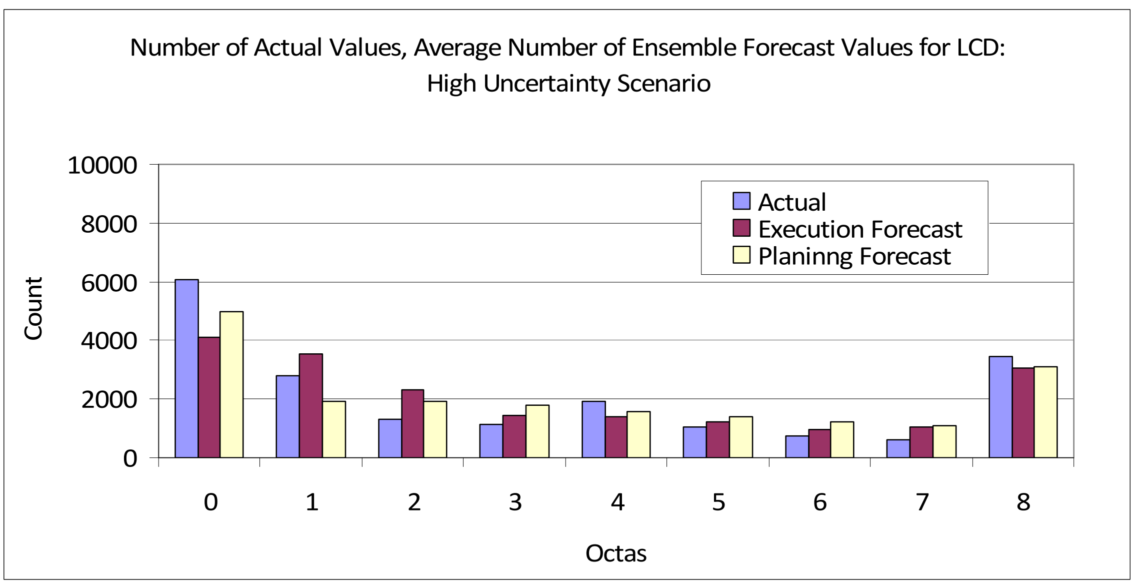

Figure 65. Number of Actual Values and Average Number of Ensemble Forecast Values for LCD: High Uncertainty Scenario. Values represent the distribution after applying the fix for WIAT's truncation of forecast values.

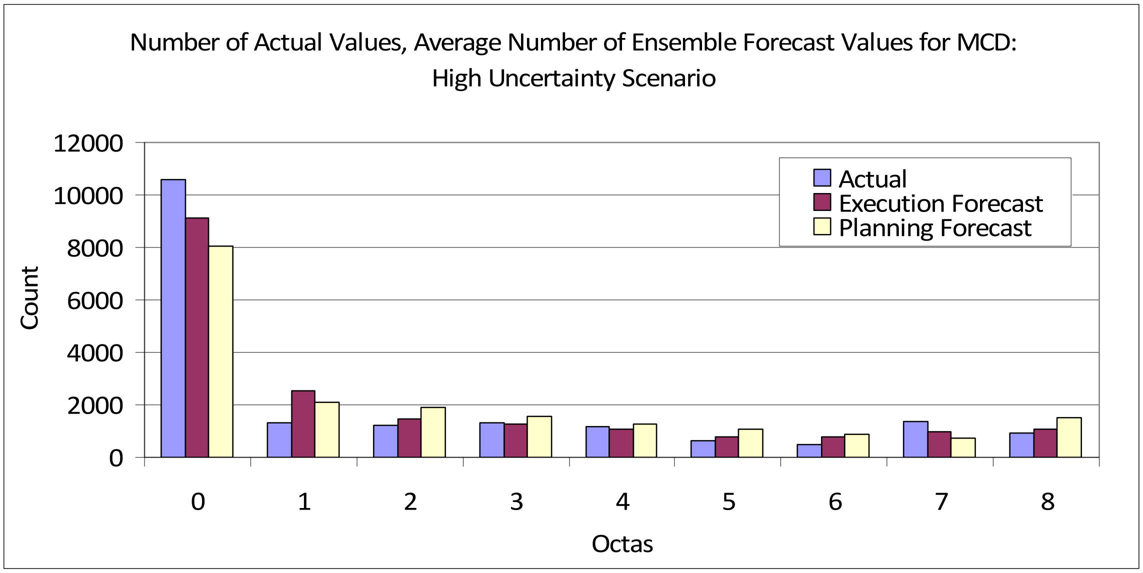

Figure 66. Number of Actual Values and Average Number of Ensemble Forecast Values for MCD: High Uncertainty Scenario. Values represent the distribution prior to applying the fix for WIAT's truncation of forecast values.

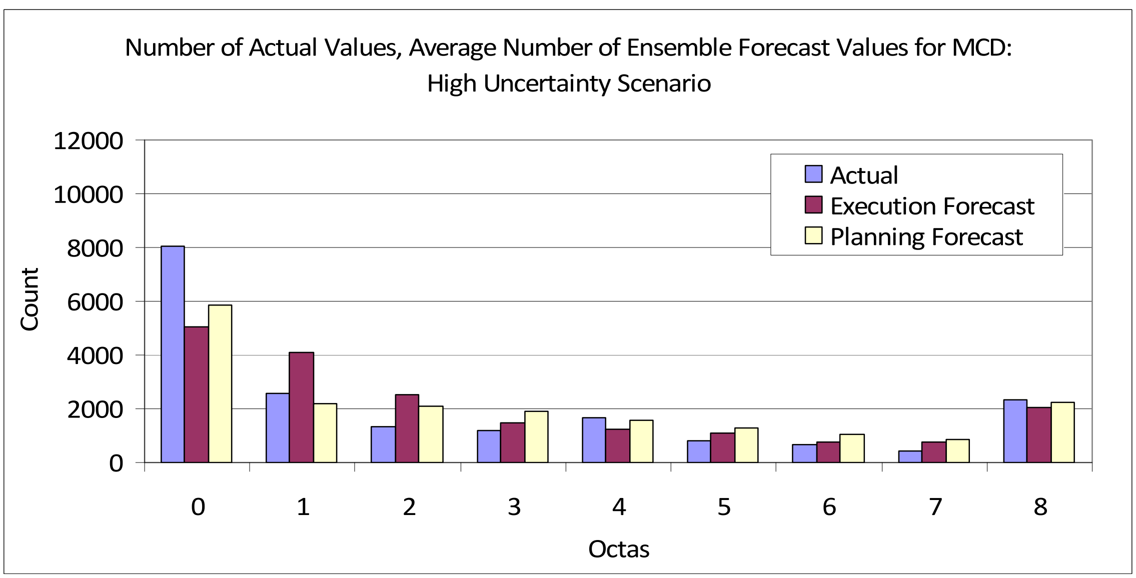

Figure 67. Number of Actual Values and Average Number of Ensemble Forecast Values for MCD: High Uncertainty Scenario. Values represent the distribution after applying the fix for WIAT's truncation of forecast values.

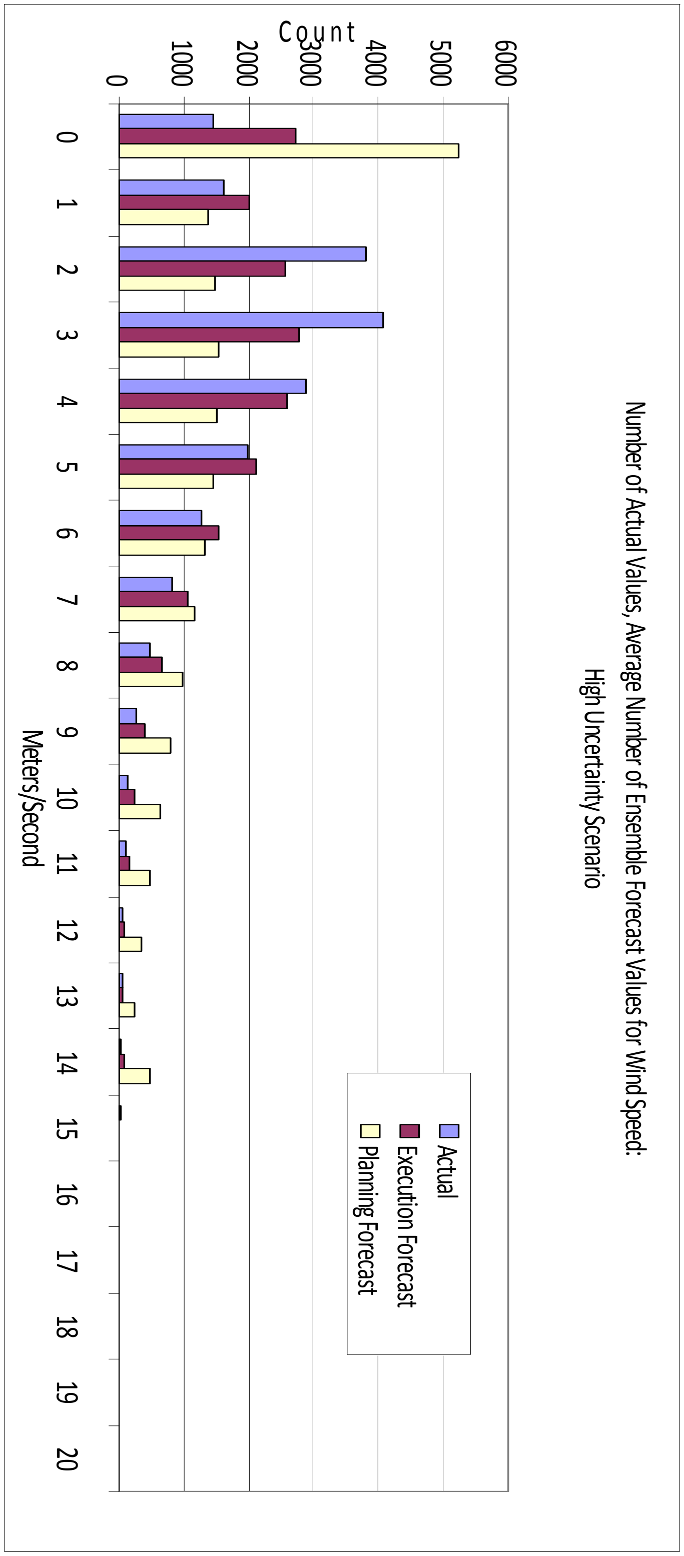

Figure 68. Number of Actual Values and Average Number of Ensemble Forecast Values for Wind Speed: High Uncertainty Scenario. Also Figure 18.

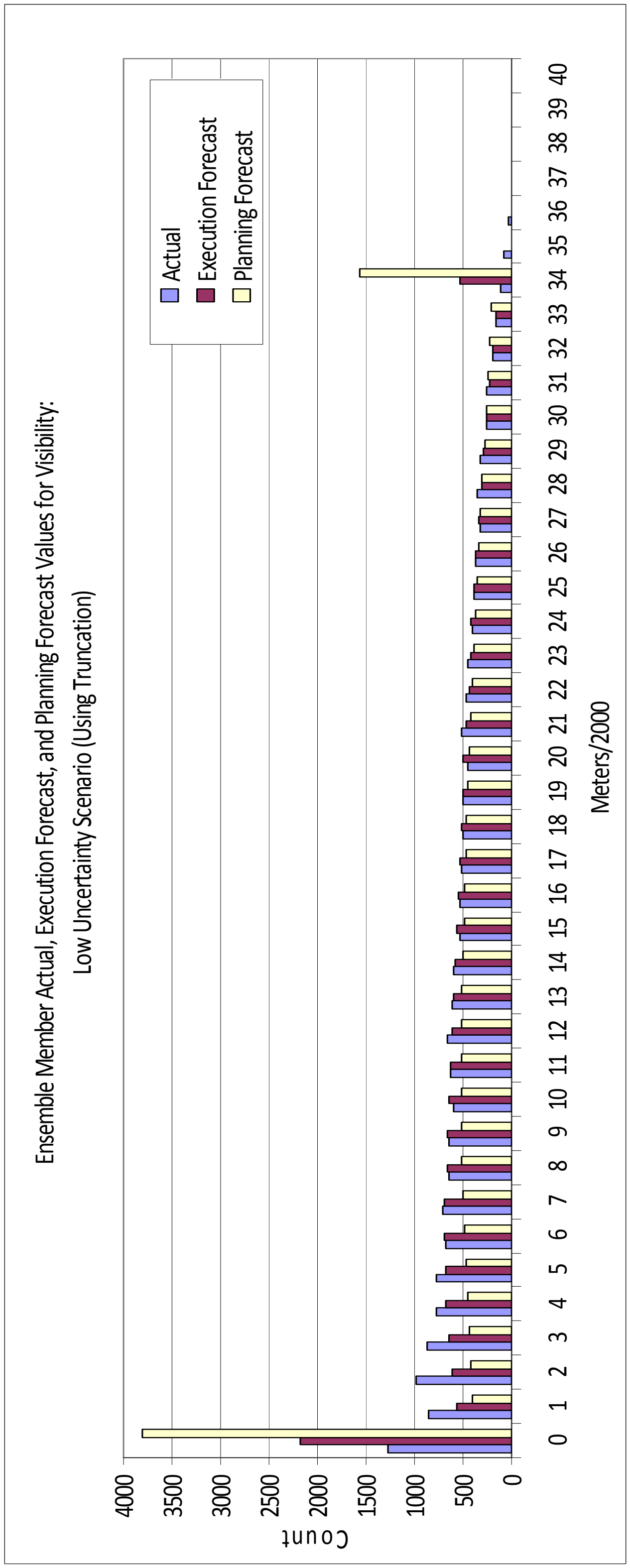

Figure 69. Number of Actual Values and Average Number of Ensemble Forecast Values for Visibility: High Uncertainty Scenario.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E: RELIABILITY DIAGRAMS

This appendix shows reliability diagrams for low- and mid-level cloud cover, wind speed, and visibility for each of the low and high uncertainty scenarios.

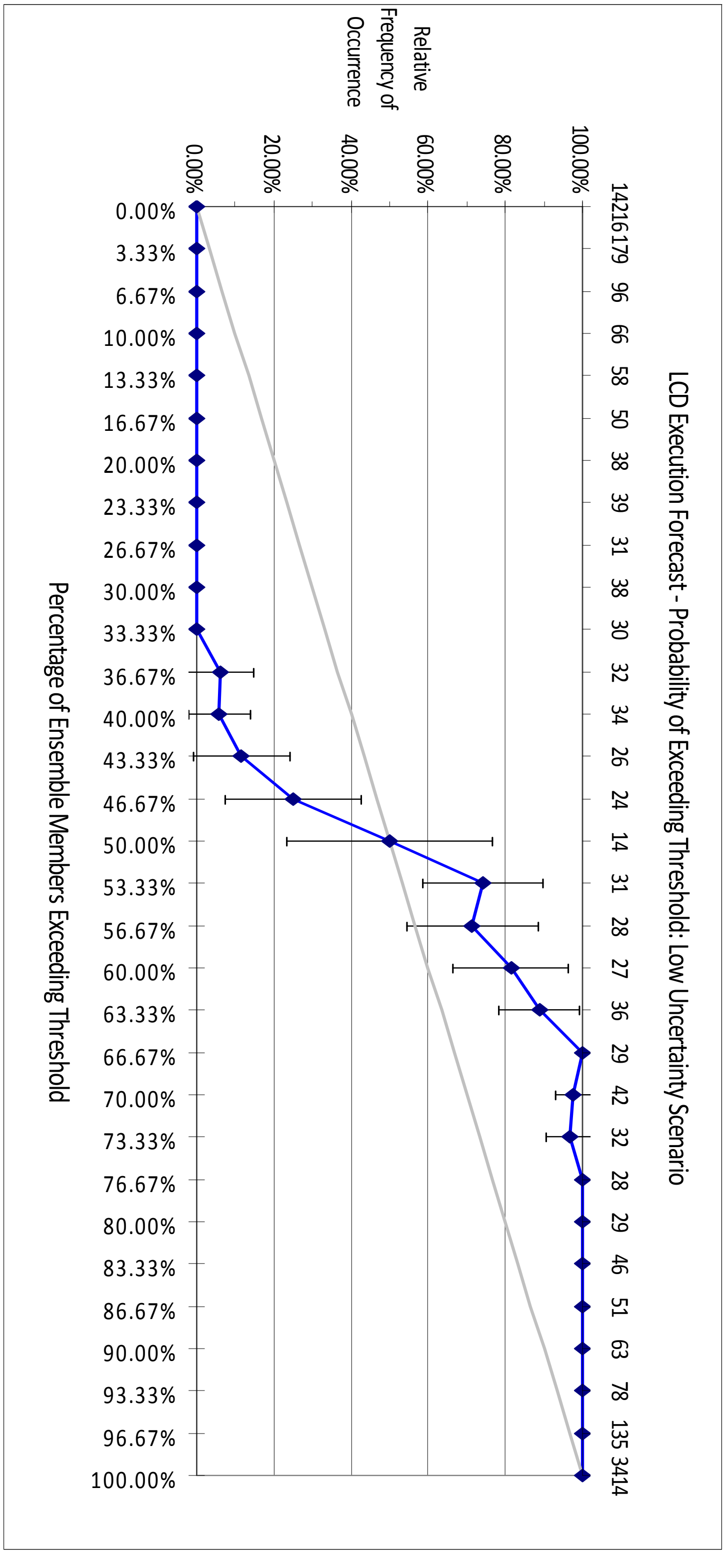

Figure 70. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for LCD Execution Forecast, Low Uncertainty Scenario. Also Figure 21.

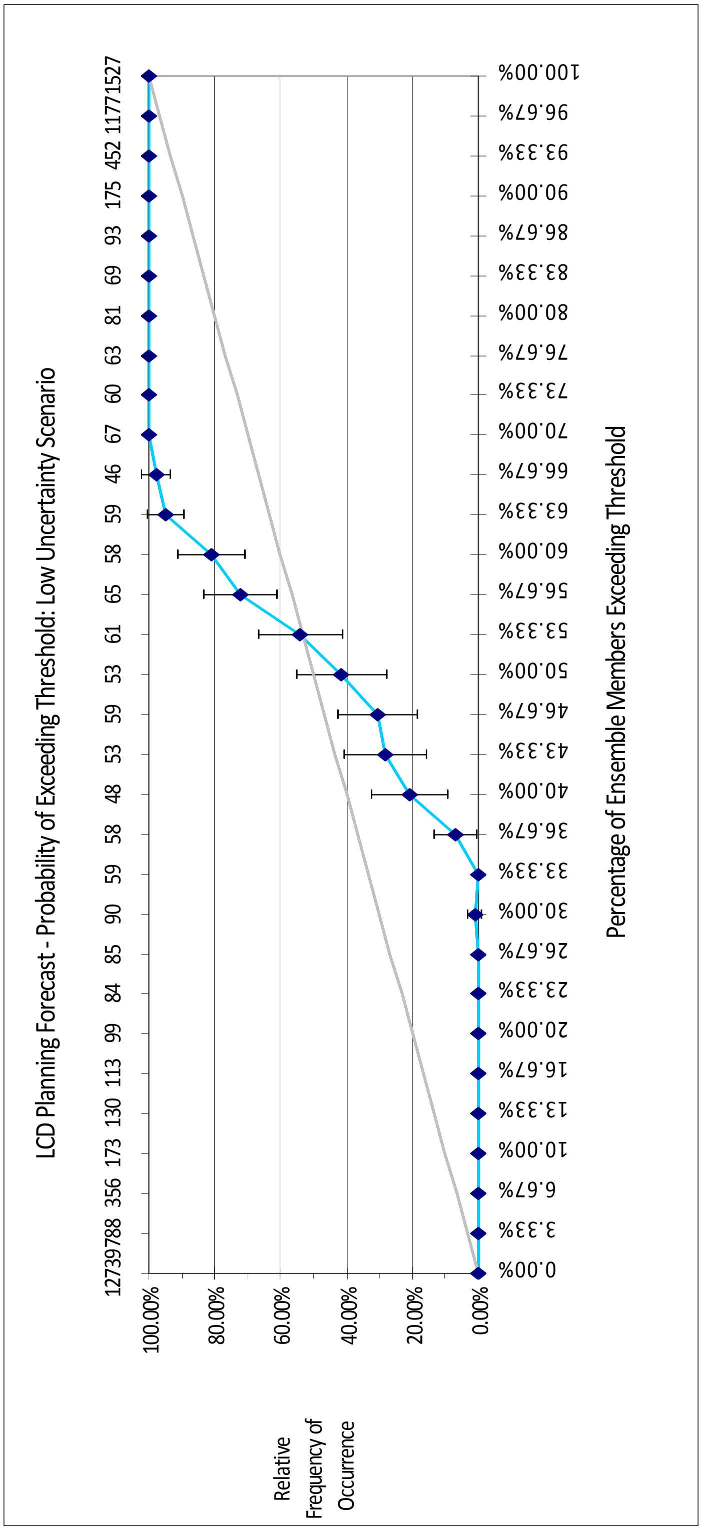

Figure 71. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for LCD Planning Forecast, Low Uncertainty Scenario.

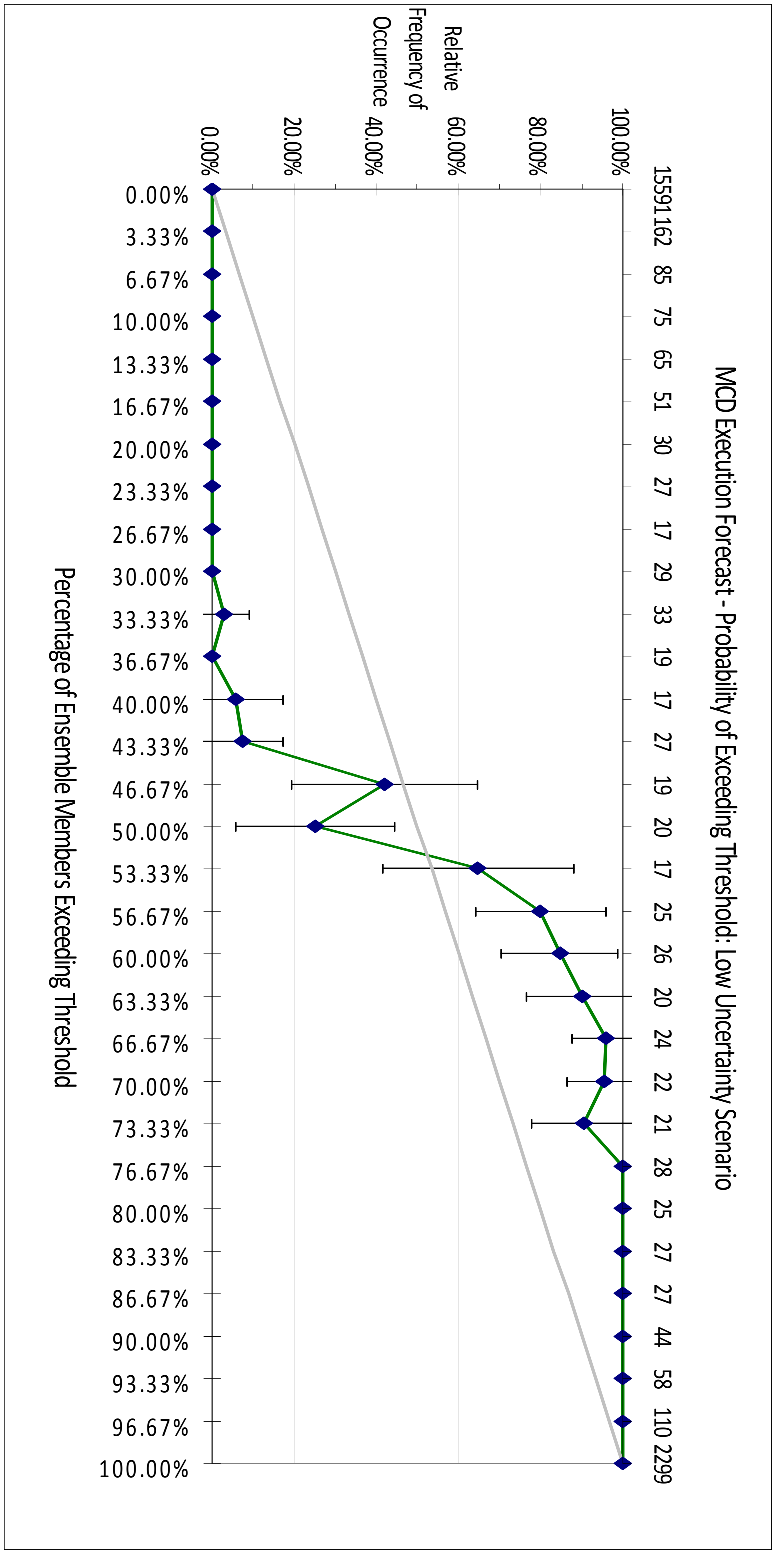

Figure 72. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for MCD Execution Forecast, Low Uncertainty Scenario.

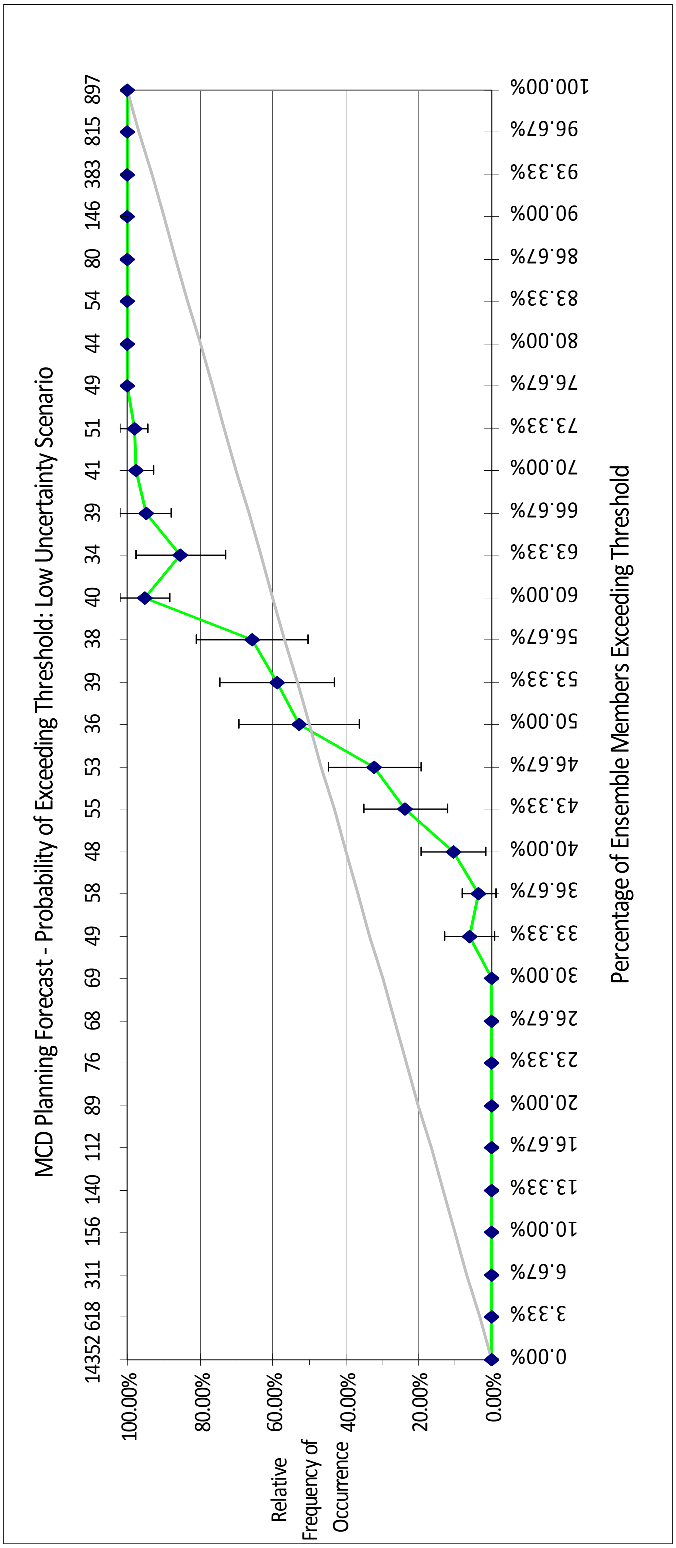

Figure 73. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for MCD Planning Forecast, Low Uncertainty Scenario.

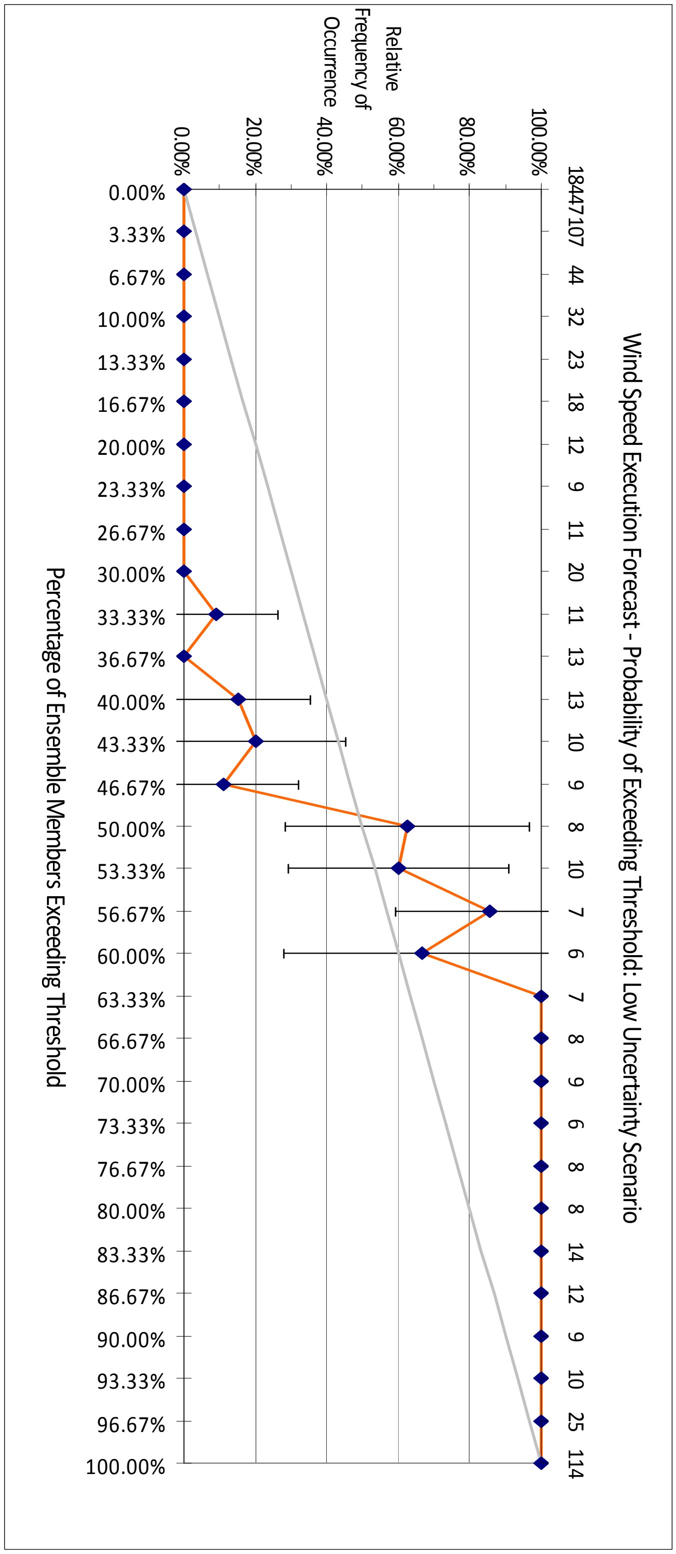

Figure 74. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for Wind Speed Execution Forecast, Low Uncertainty Scenario. Also Figure 22.

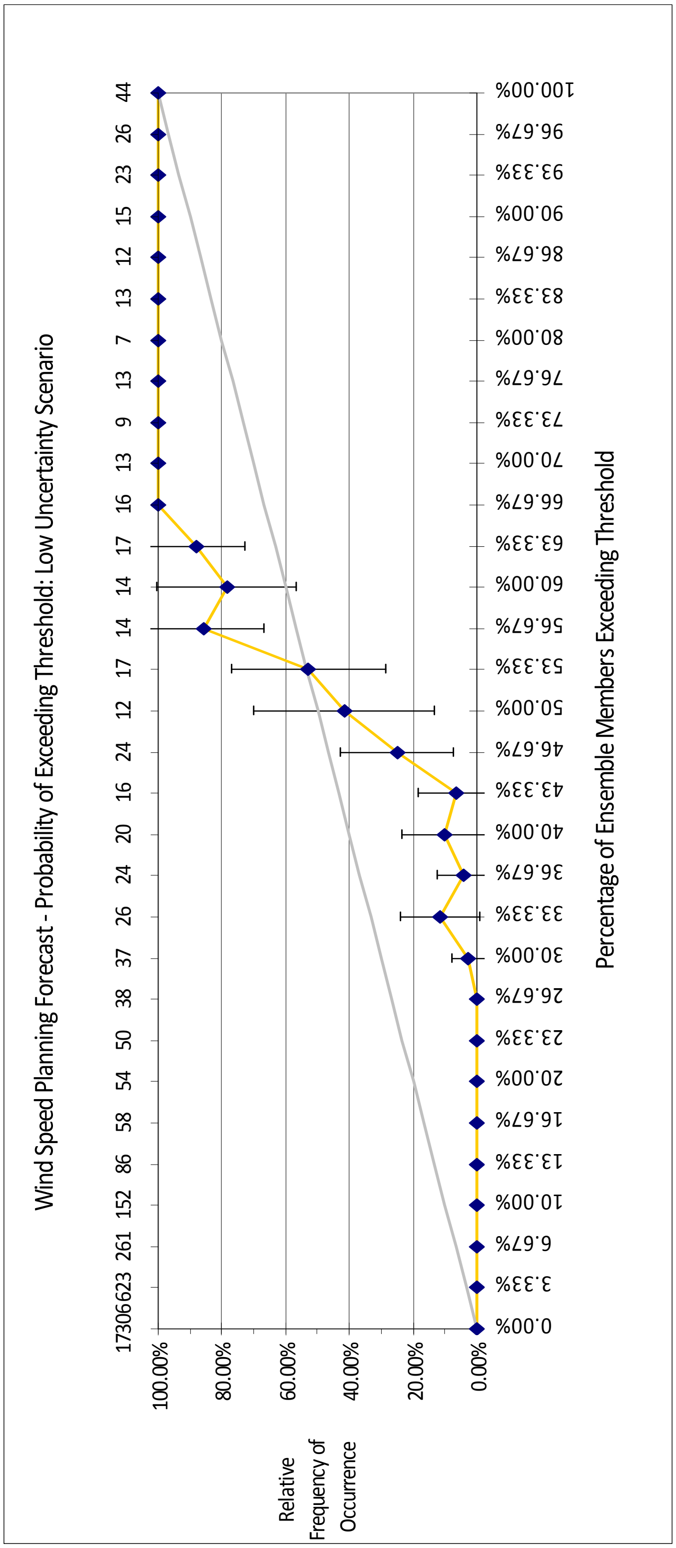

Figure 75. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for Wind Speed Planning Forecast, Low Uncertainty Scenario.

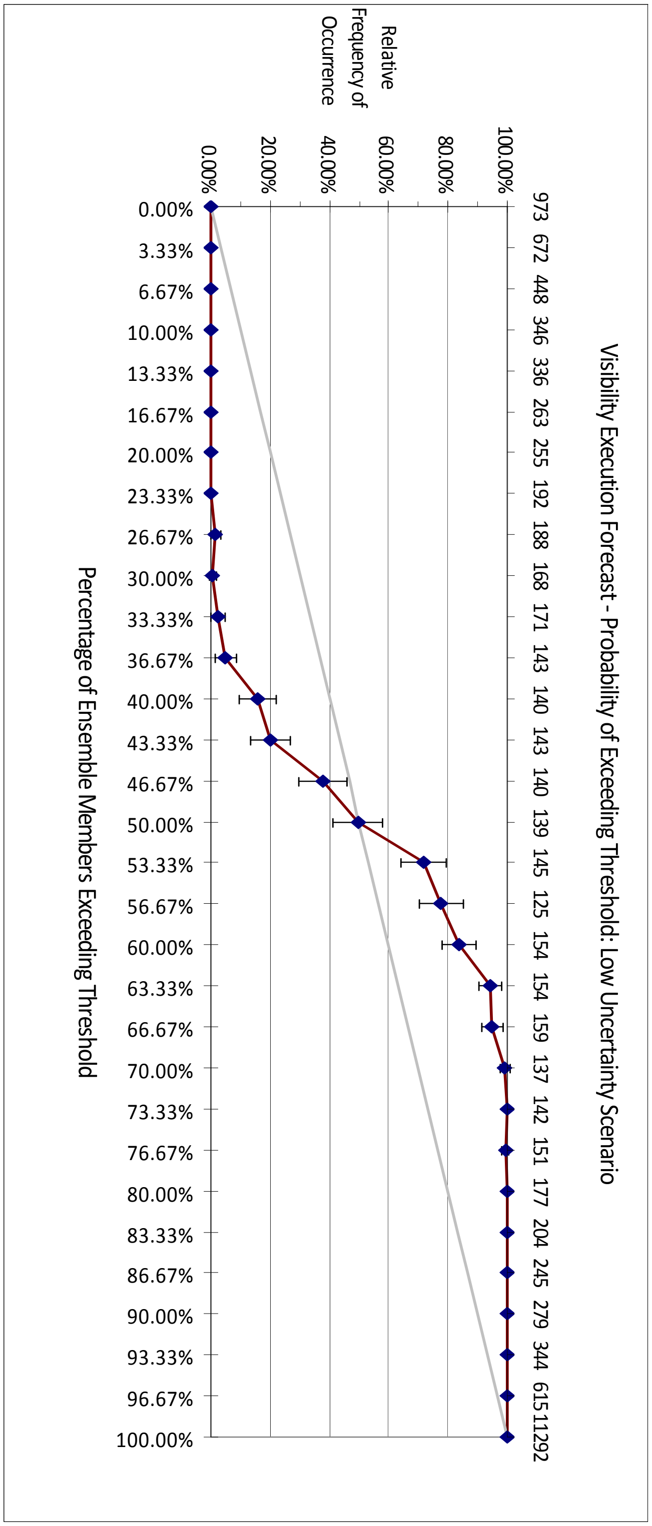

Figure 76. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for Visibility Execution Forecast, Low Uncertainty Scenario.

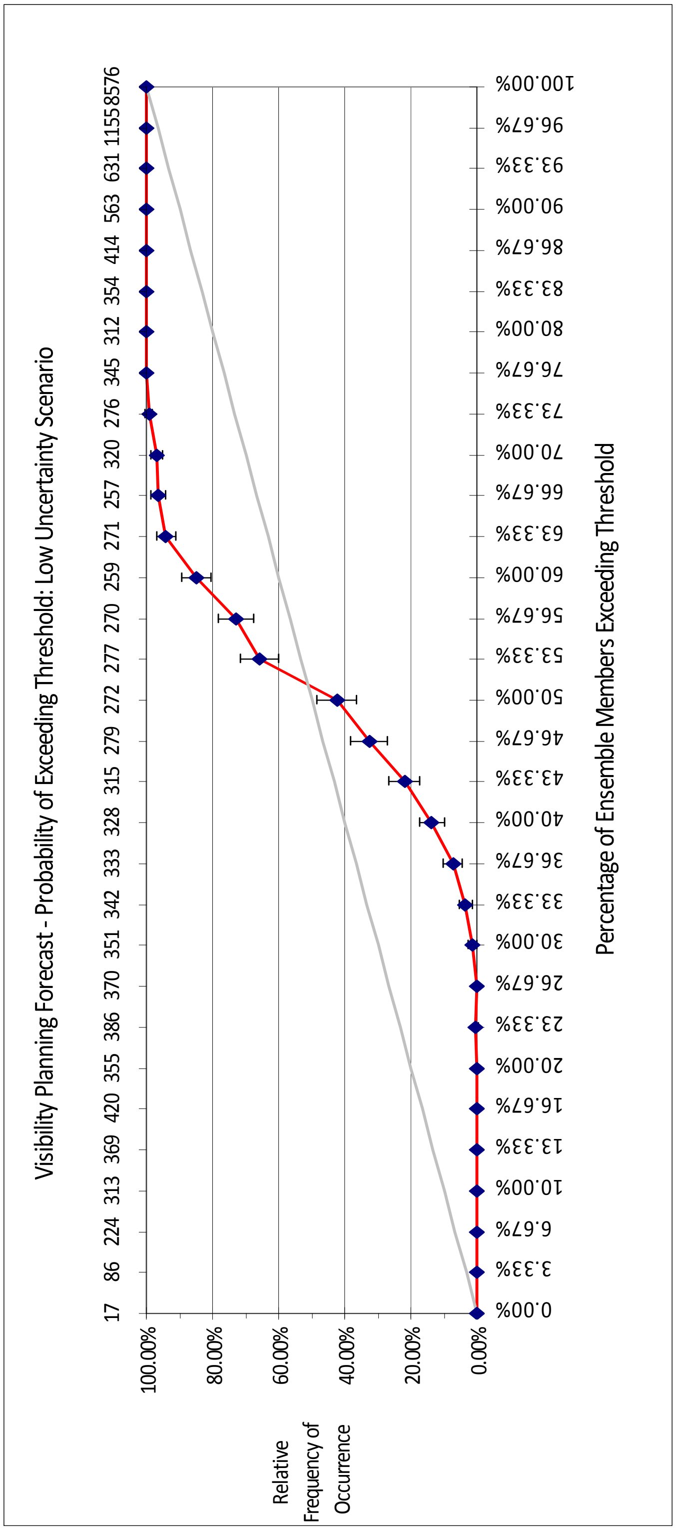

Figure 77. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for Visibility Planning Forecast, Low Uncertainty Scenario.

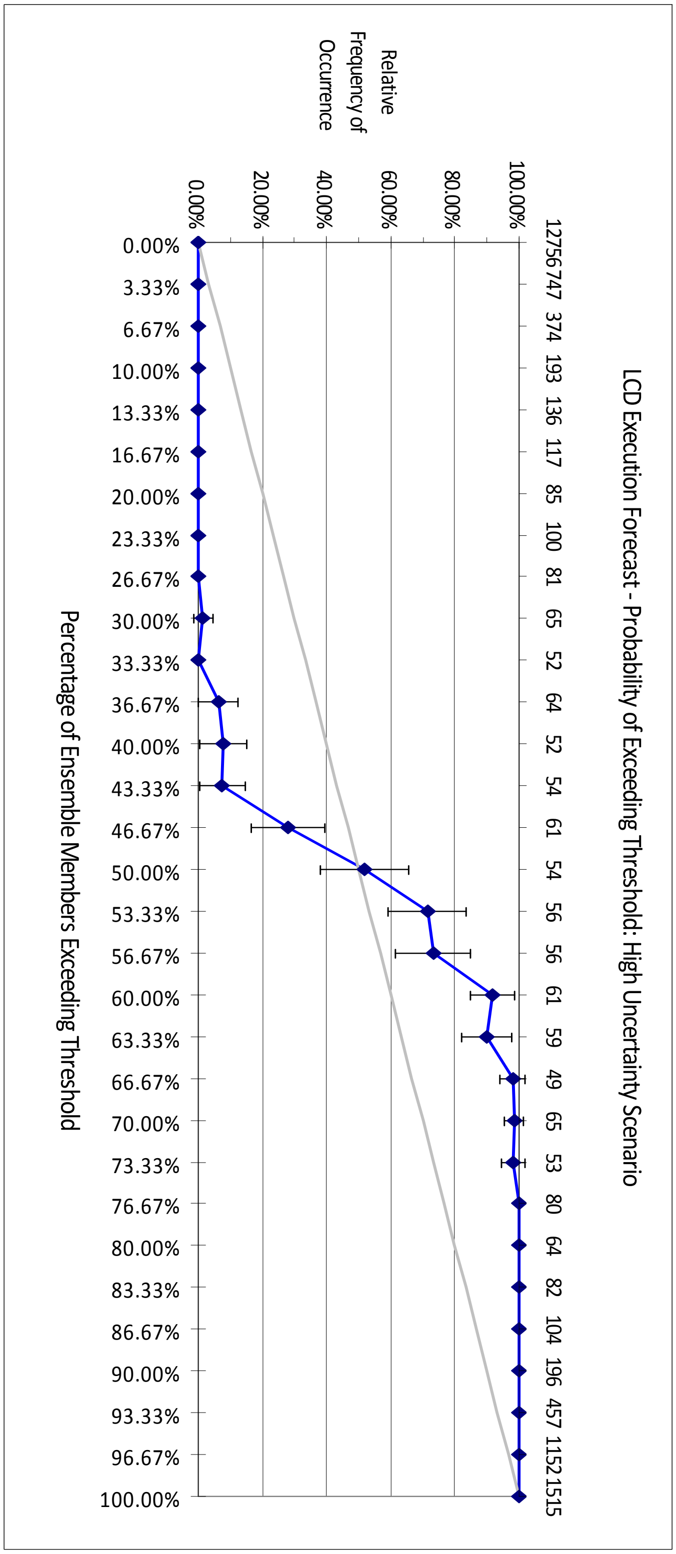

Figure 78. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for LCD Execution Forecast, High Uncertainty Scenario.

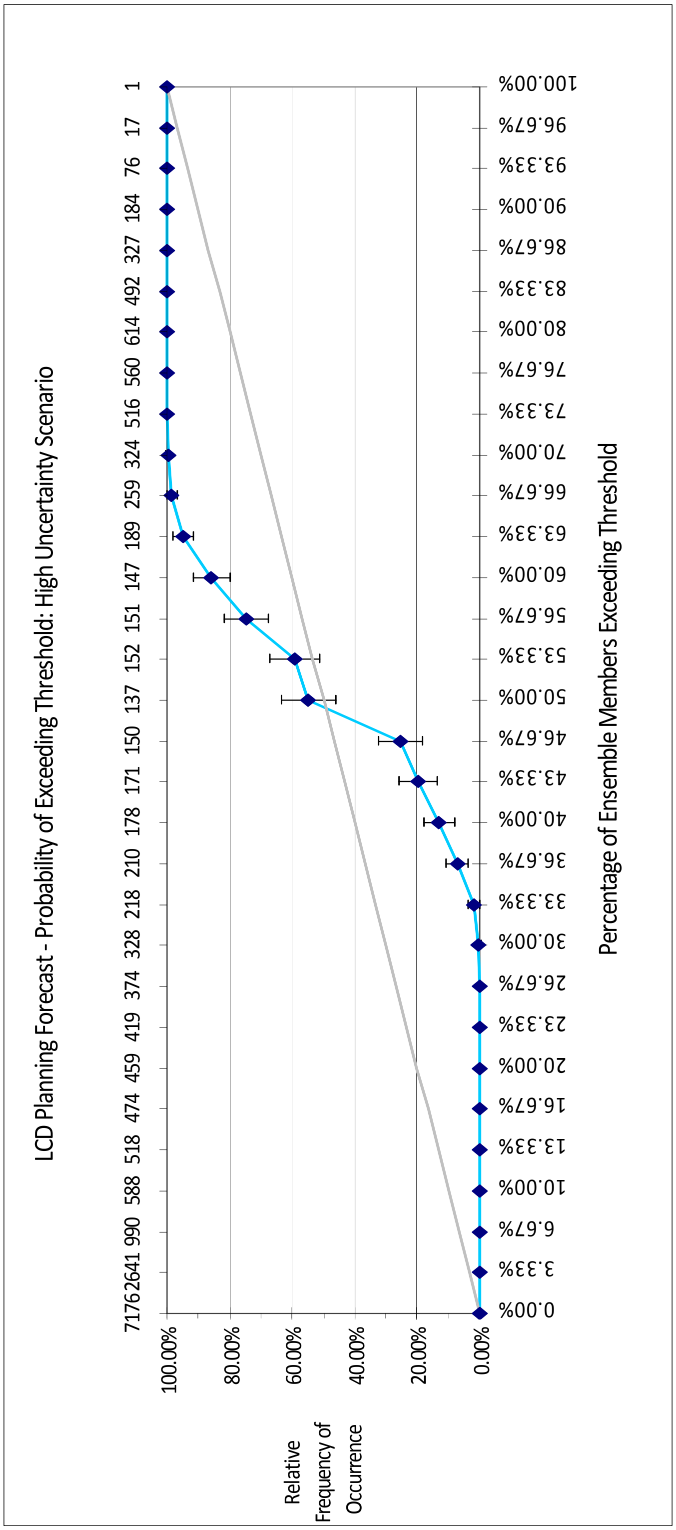

Figure 79. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for LCD Planning Forecast, High Uncertainty Scenario.

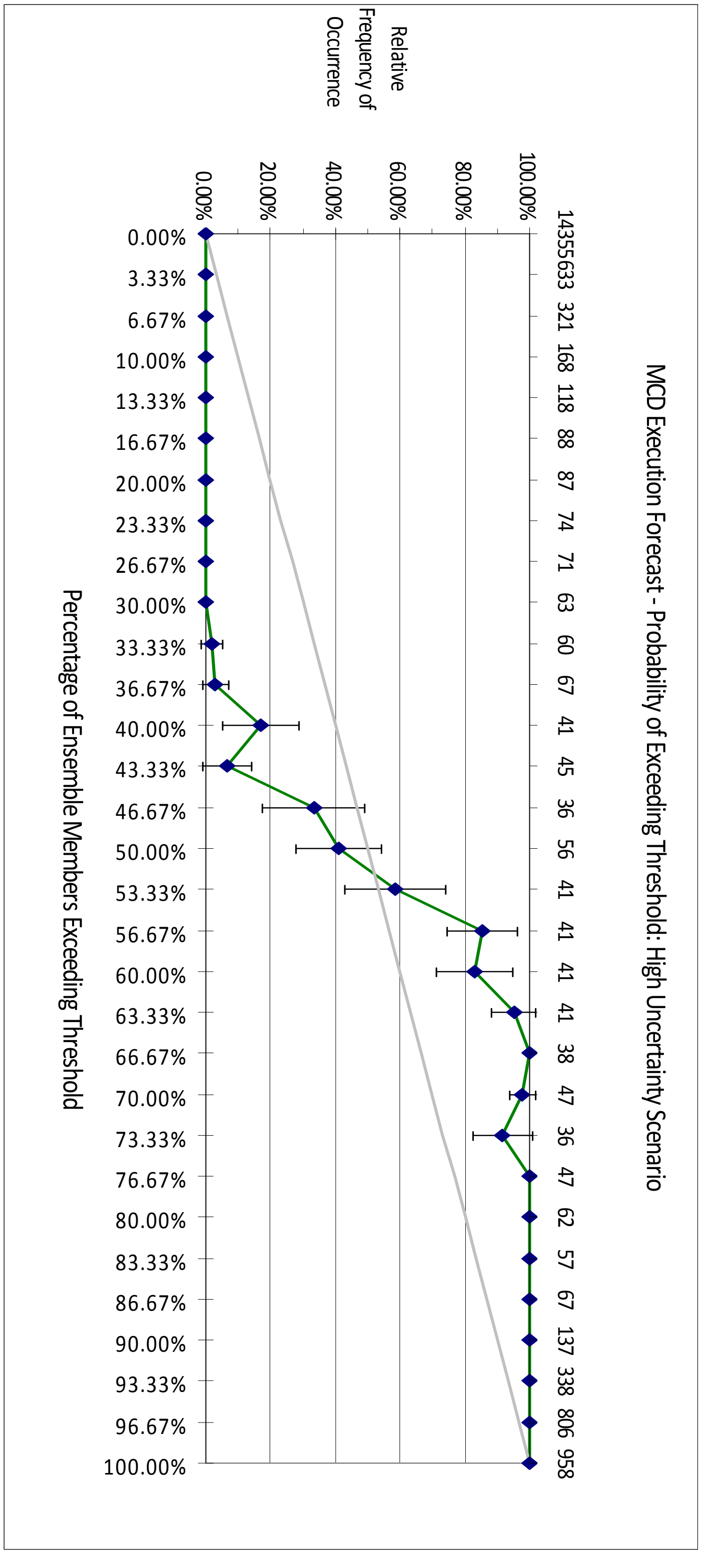

Figure 80. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for MCD Execution Forecast, High Uncertainty Scenario.

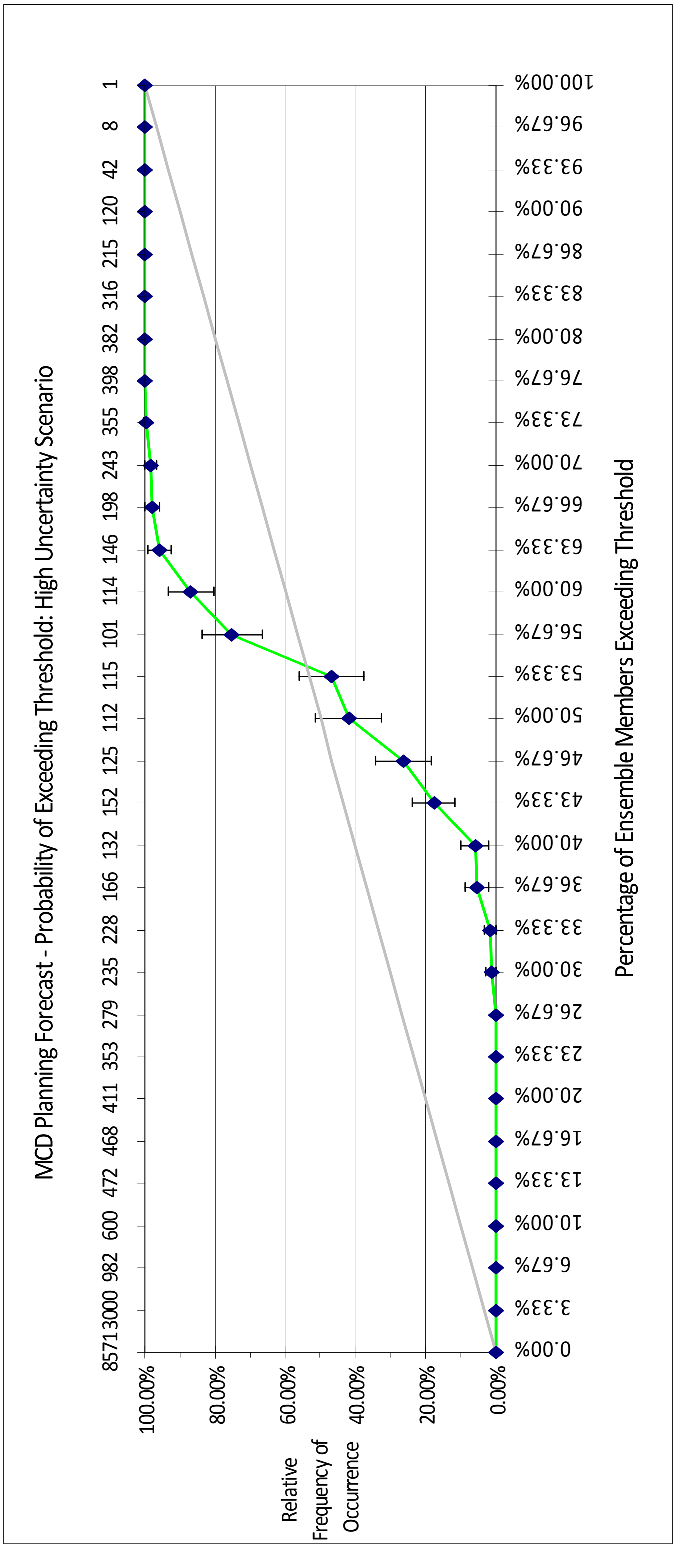

Figure 81. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for MCD Planning Forecast, High Uncertainty Scenario.

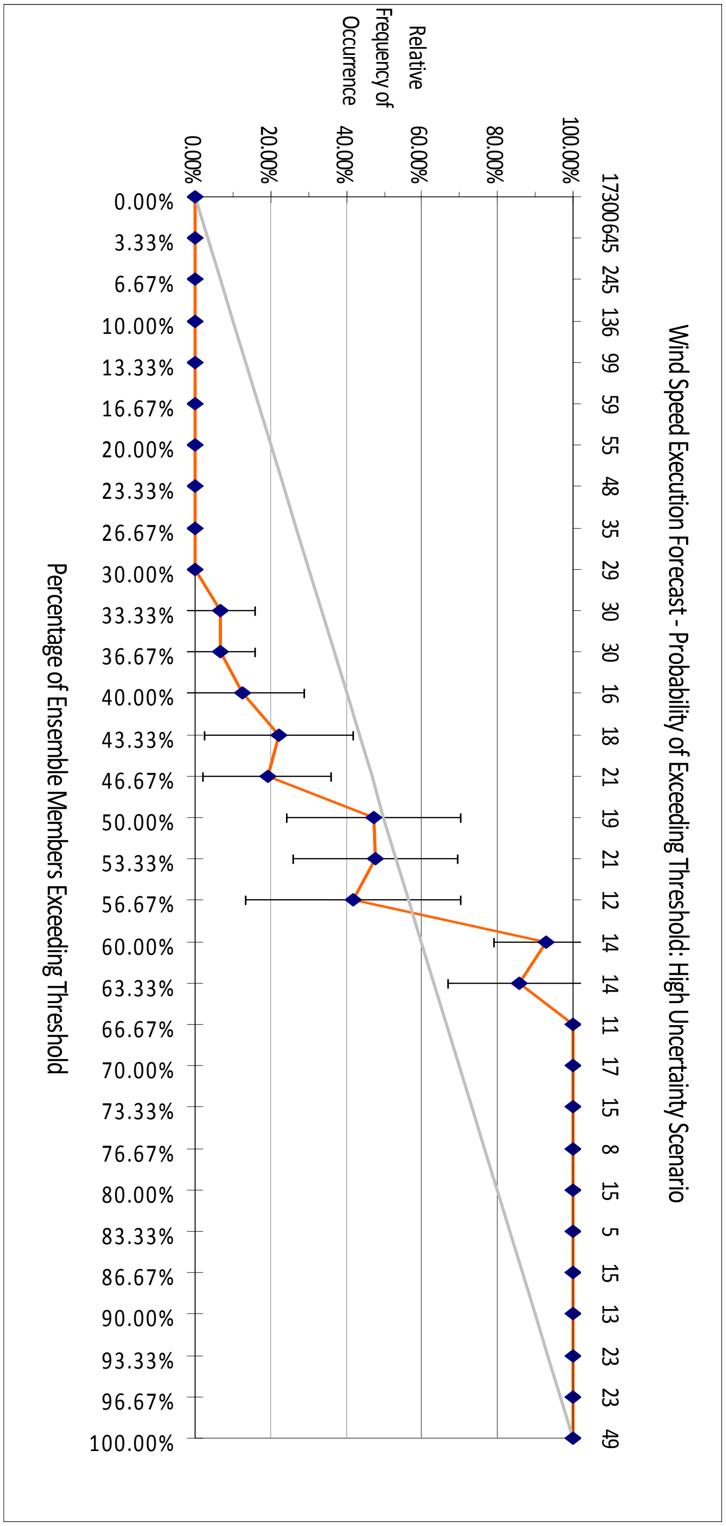

Figure 82. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for Wind Speed Execution Forecast, High Uncertainty Scenario.

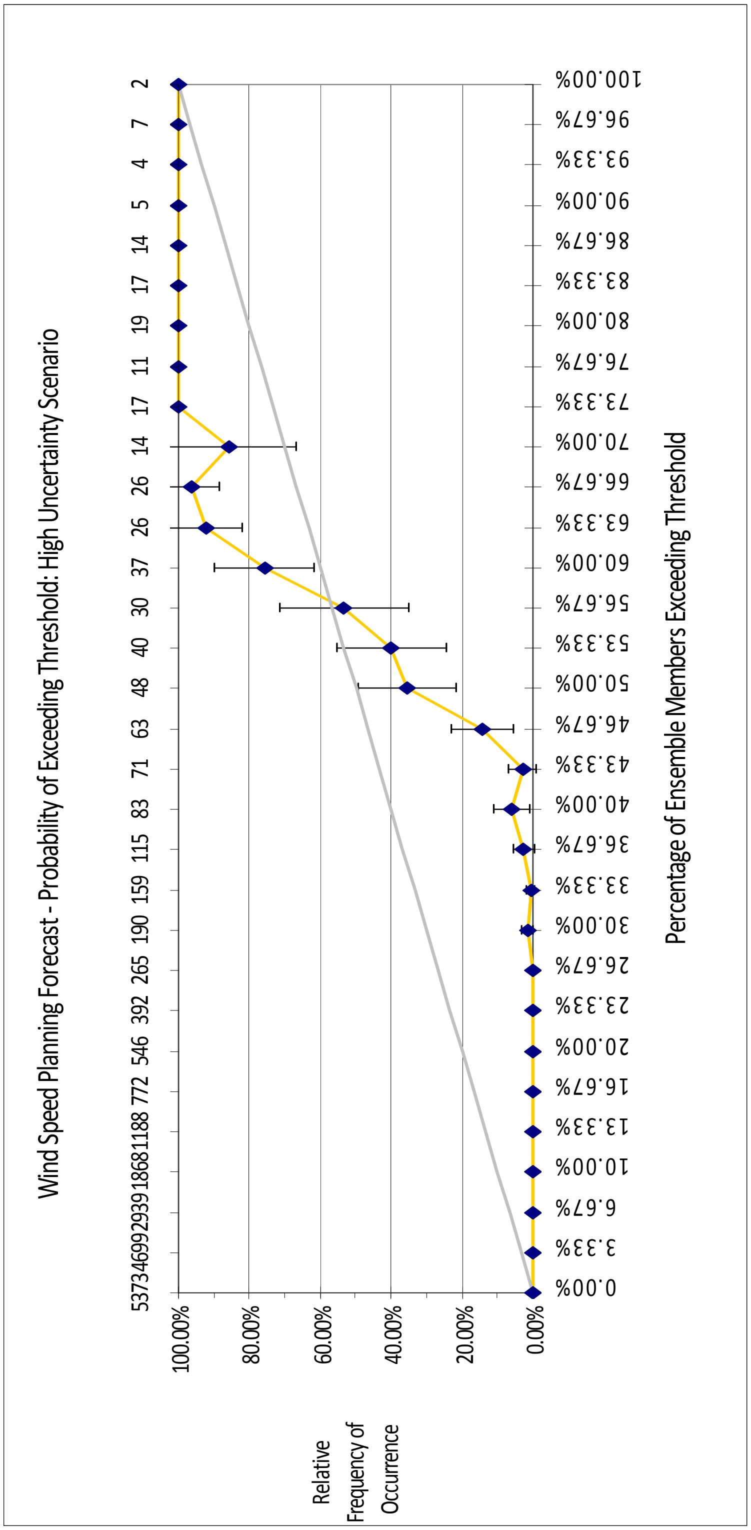

Figure 83. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for Wind Speed Planning Forecast, High Uncertainty Scenario.

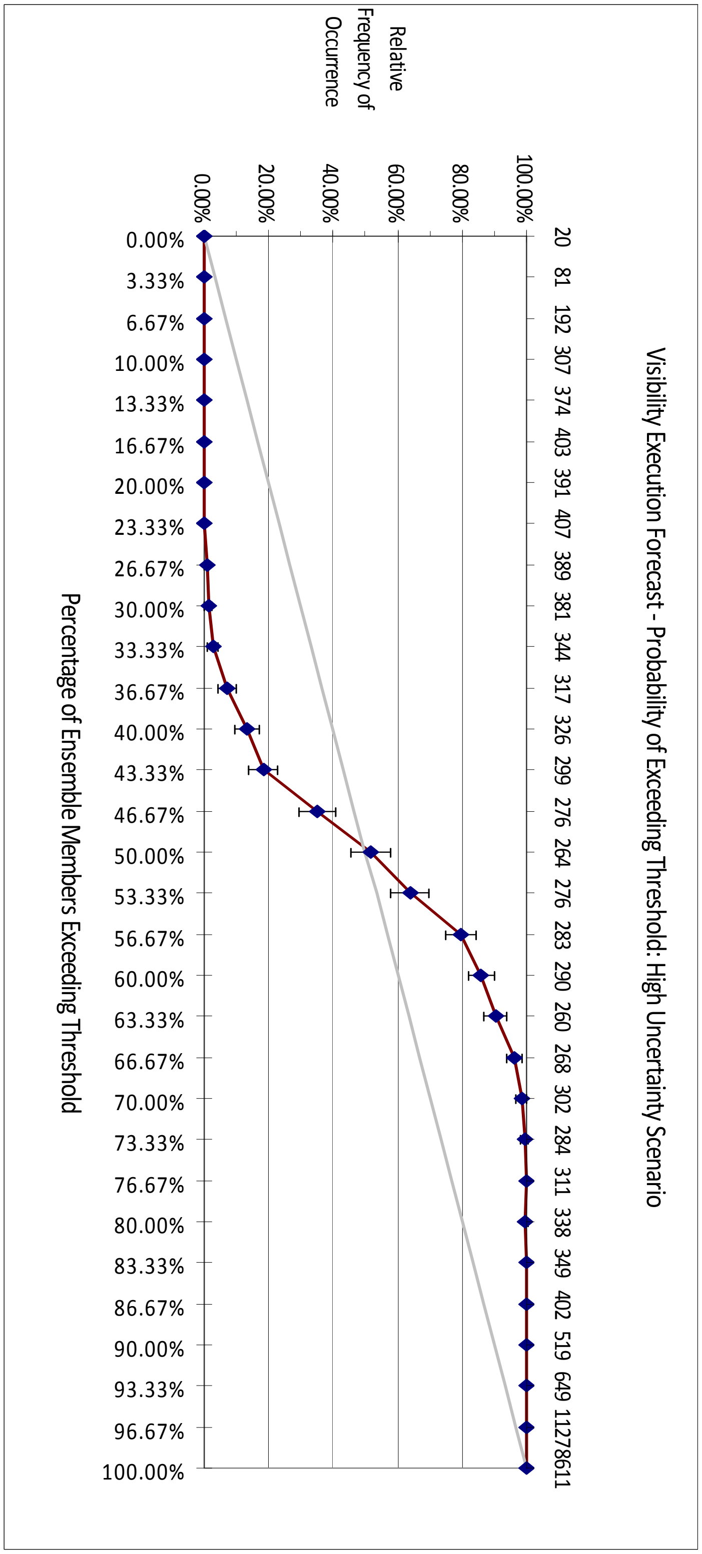

Figure 84. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for Visibility Execution Forecast, High Uncertainty Scenario.

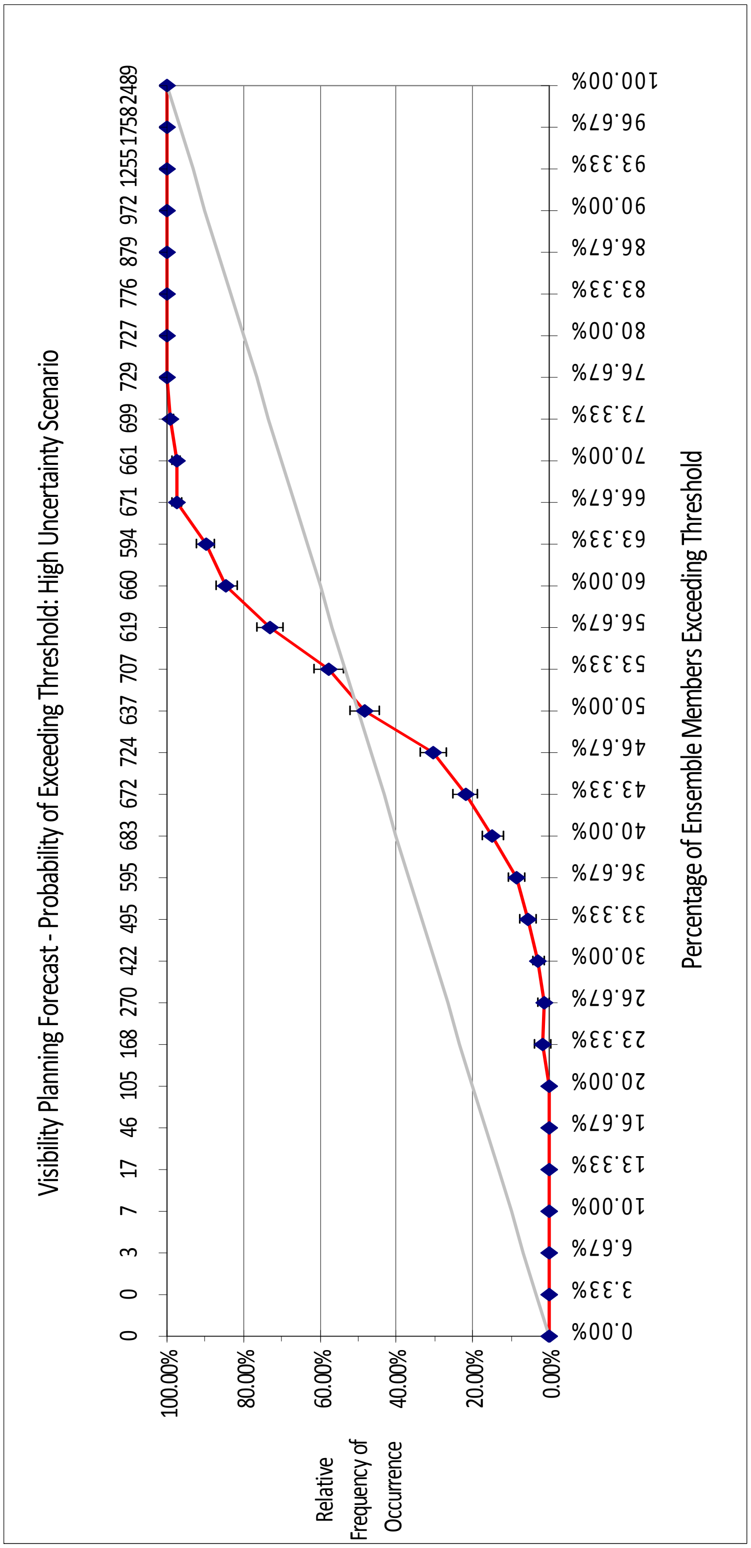

Figure 85. Probability of Exceeding Threshold vs. Relative Frequency of Occurrence for Visibility Planning Forecast, High Uncertainty Scenario.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX F: OPERATIONAL IMPACT METRICS FOR WIAT*

This appendix shows average operational impact metrics for each probability threshold in the low and high uncertainty scenarios.

Operational Impact Metrics for Stochastic Low and High Uncertainty Scenarios (Uncalibrated)																
Ensemble Execution Sigma	Ensemble Planning Sigma	Probability Threshold	Total Blue Missions Impacted	Total Blue Mission Impacts	Total Missions with 25%+ Increase in Flight Route Distance	Total Missions Unable to Conduct BDA	AvgWeaponPk	Number of KI / CAS Kills	Number of STK Kills	Number of Downed Blue Manned Aircraft	Number of Downed Blue UAVs	Number of STK Missions	Number of KI Missions	Number of CAS Missions	Strike kills per 100 Missions	KICAS kills per 100 Missions
0.5	1	12.5%	120.20	144.37	48.87	95.50	0.36599	33.40	62.10	1.37	0.33	91.13	46.50	100.00	68.14	22.80
0.5	1	25.0%	127.50	152.80	51.83	100.97	0.36998	35.80	64.70	1.40	0.40	94.57	48.83	100.00	68.42	24.05
0.5	1	37.5%	142.20	167.57	58.87	108.70	0.36520	38.33	66.67	1.63	0.30	101.53	54.73	100.00	65.66	24.77
0.5	1	50.0%	148.73	171.10	60.10	111.00	0.36844	40.07	67.50	2.13	0.27	105.33	55.83	100.00	64.08	25.71
0.5	1	62.5%	151.50	176.37	60.83	115.53	0.36087	40.17	68.33	1.37	0.27	107.87	58.57	100.00	63.35	25.33
0.5	1	75.0%	154.30	177.57	61.03	116.53	0.36451	41.57	68.77	2.20	0.23	108.97	62.03	100.00	63.11	25.65
0.5	1	87.5%	155.07	179.17	61.93	117.23	0.36400	41.37	69.37	1.97	0.37	109.70	63.20	100.00	63.23	25.35
0.5	1	100.0%	165.10	205.40	73.83	131.57	0.36156	48.20	72.10	1.20	0.33	110.00	71.90	100.00	65.55	28.04
1	2.5	12.5%	109.13	129.27	43.13	86.13	0.37157	30.97	59.00	1.07	0.30	85.17	43.27	100.00	69.28	21.61
1	2.5	25.0%	120.33	144.70	48.23	96.47	0.36843	34.60	62.27	0.77	0.57	90.50	47.00	100.00	68.80	23.54
1	2.5	37.5%	131.90	156.60	54.17	102.43	0.37414	37.70	65.10	1.63	0.27	97.07	50.63	100.00	67.07	25.03
1	2.5	50.0%	147.63	172.20	61.03	111.17	0.36722	40.07	67.27	1.10	0.30	104.37	56.40	100.00	64.45	25.62
1	2.5	62.5%	152.97	178.50	61.70	116.80	0.35956	40.70	68.40	1.83	0.17	108.67	60.30	100.00	62.94	25.39
1	2.5	75.0%	156.87	183.93	62.97	120.97	0.35886	44.10	67.73	1.60	0.23	109.73	63.80	100.00	61.73	26.92
1	2.5	87.5%	158.67	190.20	65.53	124.67	0.35972	46.40	68.33	1.93	0.17	110.00	67.77	100.00	62.12	27.66
1	2.5	100.0%	165.10	205.40	73.83	131.57	0.36156	48.20	72.10	1.20	0.33	110.00	71.90	100.00	65.55	28.04

Table 11. Operational Impact Metrics for Uncalibrated Low and High Uncertainty Scenarios.

Operational Impact Metrics for Stochastic Low and High Uncertainty Scenarios (Calibrated)																
Ensemble Execution Sigma	Ensemble Planning Sigma	Probability Threshold	Total Blue Missions Impacted	Total Blue Mission Impacts	Total Missions with 25%+ Increase in Flight Route Distance	Total Missions Unable to Conduct BDA	AvgWeaponPk	Number of KI / CAS Kills	Number of STK Kills	Number of Downed Blue Manned Aircraft	Number of Downed Blue UAVs	Number of STK Missions	Number of KI Missions	Number of CAS Missions	Strike kills per 100 Missions	KICAS kills per 100 Missions
0.5	1	12.5%	130.63	151.90	51.77	100.13	0.37279	36.60	64.13	1.10	0.23	95.67	50.03	100.00	67.04	24.39
0.5	1	25.0%	138.50	160.37	54.63	105.73	0.37067	38.83	65.43	1.27	0.23	98.63	52.73	100.00	66.34	25.43
0.5	1	37.5%	143.40	166.03	58.10	107.93	0.36585	37.67	66.77	1.93	0.27	102.40	53.70	100.00	65.20	24.51
0.5	1	50.0%	148.43	171.90	59.80	112.10	0.36531	39.63	67.73	2.13	0.30	105.53	56.87	100.00	64.18	25.27
0.5	1	62.5%	152.50	177.10	61.23	115.87	0.36681	42.37	68.30	1.97	0.27	107.53	58.73	100.00	63.52	26.69
0.5	1	75.0%	154.40	176.37	60.47	115.90	0.37019	42.70	68.93	1.30	0.10	108.73	60.60	100.00	63.40	26.59
0.5	1	87.5%	155.67	178.40	60.77	117.63	0.36605	42.70	68.93	1.90	0.33	109.13	61.20	100.00	63.16	26.49
0.5	1	100.0%	165.10	205.40	73.83	131.57	0.36156	48.20	72.10	1.20	0.33	110.00	71.90	100.00	65.55	28.04
1	2.5	12.5%	126.77	151.07	50.70	100.37	0.36808	36.17	63.47	1.27	0.33	92.90	49.13	100.00	68.32	24.25
1	2.5	25.0%	134.13	155.90	53.57	102.33	0.37413	37.57	65.13	1.43	0.27	97.23	51.67	100.00	66.99	24.77
1	2.5	37.5%	142.07	166.87	58.20	108.67	0.36536	38.43	66.47	1.33	0.27	101.30	53.53	100.00	65.61	25.03
1	2.5	50.0%	147.93	172.13	60.17	111.97	0.36753	40.53	67.50	2.07	0.20	105.23	56.10	100.00	64.14	25.97
1	2.5	62.5%	149.93	176.23	61.03	115.20	0.35994	40.03	68.03	1.40	0.17	107.27	58.40	100.00	63.42	25.27
1	2.5	75.0%	152.63	177.23	61.50	115.73	0.36579	41.47	68.73	1.00	0.10	108.87	60.20	100.00	63.14	25.88
1	2.5	87.5%	156.07	182.13	62.37	119.77	0.36051	42.37	69.03	1.77	0.27	109.60	62.40	100.00	62.99	26.09
1	2.5	100.0%	165.10	205.40	73.83	131.57	0.36156	48.20	72.10	1.20	0.33	110.00	71.90	100.00	65.55	28.04

Table 12. Operational Impact Metrics for Calibrated Low and High Uncertainty Scenarios.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Brier, G. W., 1944: Verification of a Forecaster's Confidence and The Use of Probability Statements in Weather Forecasting. Research Report No. FR-16 for the U.S. Dept. of Commerce Weather Bureau.

Eckel, F. A., 200-8: *Introduction to Ensemble Forecasting.* Draft July 7, 2008. Naval Postgraduate School. 97 pp.

Eckel, F. A., J. G. Cunningham, and D. E. Hetke, 2008: Weather and the Calculated Risk: Exploiting Forecast Uncertainty for Operational Risk Management. *Air & Space Power Journal*, Spring 2008, 71–82.

Earth System Research Laboratory Daily Mean Composites Web site, http://www.esrl.noaa.gov/psd/data/composites/day/ [accessed May 12, 2010].

Fleet Numerical Meteorology and Oceanography Web site, https://www.fnmoc.navy.mil/efs/dynamic/US058VMET-GIFwxg.EFS.no_amer_prob_precip_96.gif [accessed September 9, 2008].

Kalnay, E., 2004: Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press, 341 pp.

Lorenz, E. N., 1963: Deterministic Nonperiodic Flow. *J. Atmos. Sci.*, **20**, 130–141.

Morone, Lauren L. and S. J. Pond, 1994. *The Rapid Update Cycle Part II: Verification and Evaluation.* National Weather Service Office of Meteorology Technical Procedures Bulletin, Series No. 429. Web site, http://www.nws.noaa.gov/om/tpb/tpb429.pdf [accessed Janurary 19, 2010].

Murphy, A. H., 1993. What is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting. *Weather and Forecasting*, **8**, 281–293.

Sittel, M. Cloud Model Verification at the Air Force Weather Agency. Brief for Developmental Testbed Center. Web site, http://dtcenter.org/verification/presentations/Matt_Sittel_Cloud%20Verification.pdf [accessed January 22, 2010].

Systems Planning and Analysis, Inc., 2007: "The Weather Impact Assessment Tool (WIAT): A Basic Introduction to WIAT." Lecture, Naval Postgraduate School, Monterey, CA, June 11, 2007.

——. "The Weather Impact Assessment Tool (WIAT): Introduction to Metrics and Output Lab." Lecture, Naval Postgraduate School, Monterey, CA, June 11, 2007.

——. “The Weather Impact Assessment Tool (WIAT): Navigation & Input Lab.” Lecture, Naval Postgraduate School, Monterey, CA, June 11, 2007.

——. “The Weather Impact Assessment Tool (WIAT): Sample Analyses and Applications.” Lecture, Naval Postgraduate School, Monterey, CA, June 11, 2007.

Thompson, J. C., 1952: On the Operational Deficiencies in Categorical Weather Forecasts. *Bull. Amer. Meteor. Soc.* **33**, No. 6, pp. 223–226.

Titley, D. W. NAVMETOCCOMNOTE 5000, 08 January 2008, Subj: COMMANDER’S GUIDANCE.

Thompson, J. C., 1950: A Numerical Method for Forecasting Rainfall in the Los Angeles Area. *Monthly Weather Review*. **78**, No. 7, pp. 113–124.

Zhu, Y., Z. Toth, R. Wobus, D. Richardson, and K. Mylne, 2002: The Economic Value of Ensemble-Based Weather Forecasts. *Bull. Amer. Meteor. Soc*. **83**, No. 1, pp. 73–83.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, VA

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, CA

3. Rebecca Stone
   Department of Oceanography
   Naval Postgraduate School
   Monterey, CA

4. Philip Durkee
   Chair
   Department of Meteorology
   Naval Postgraduate School
   Monterey, CA

5. Luke Piepkorn
   Systems Planning & Analysis
   Sterling, VA